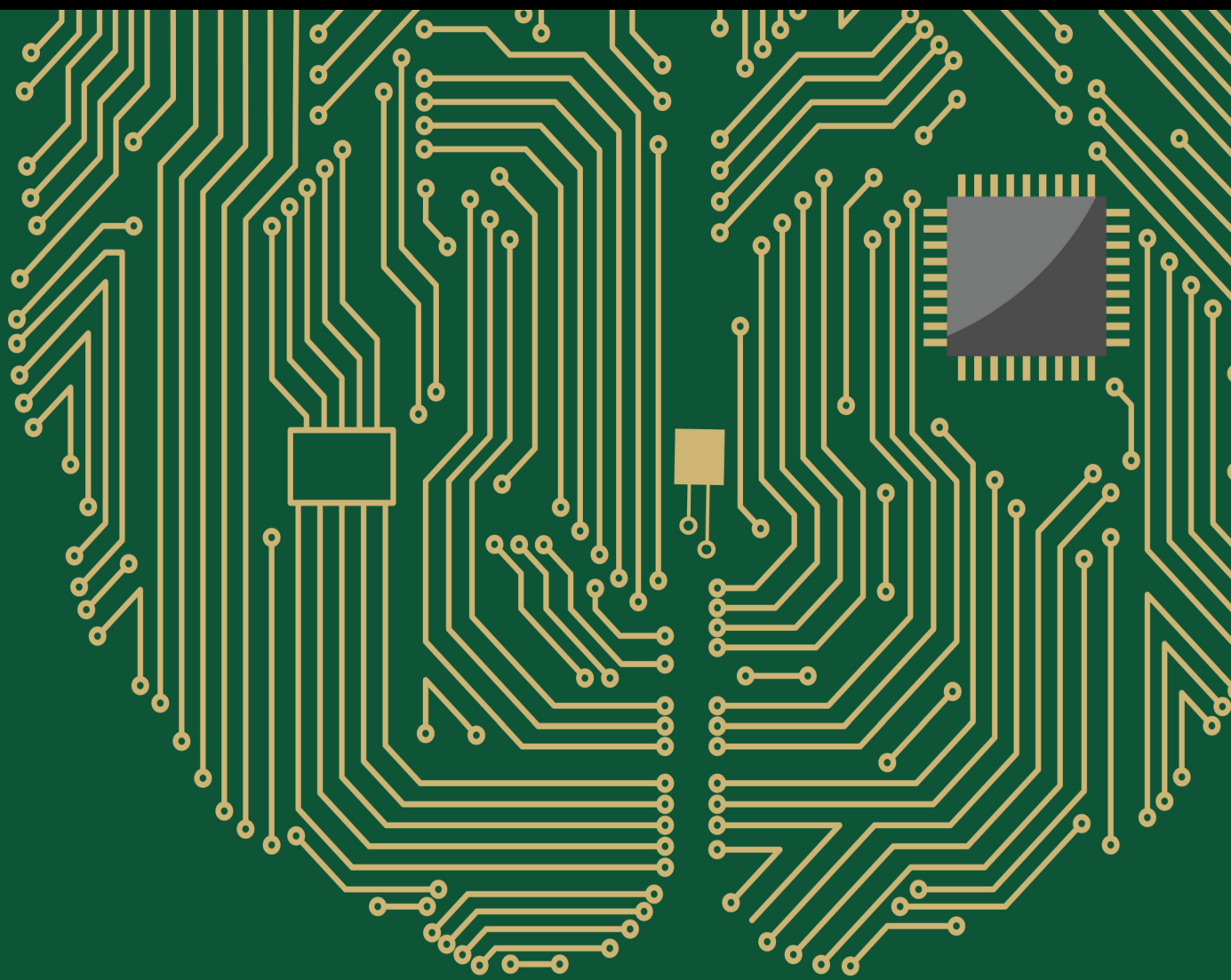


Advanced Computational Intelligence Algorithms for Signal and Image Processing

Lead Guest Editor: Baiyuan Ding

Guest Editors: Gohar Ali and Yang Gu





**Advanced Computational Intelligence
Algorithms for Signal and Image Processing**

Computational Intelligence and Neuroscience

**Advanced Computational Intelligence
Algorithms for Signal and Image
Processing**

Lead Guest Editor: Baiyuan Ding

Guest Editors: Gohar Ali and Yang Gu



Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in "Computational Intelligence and Neuroscience." 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

Andrzej Cichocki, Poland

Associate Editors

Arnaud Delorme, France
Cheng-Jian Lin , Taiwan
Saeid Sanei, United Kingdom

Academic Editors

Mohamed Abd Elaziz , Egypt
Tariq Ahanger , Saudi Arabia
Muhammad Ahmad, Pakistan
Ricardo Aler , Spain
Nouman Ali, Pakistan
Pietro Aricò , Italy
Lerina Aversano , Italy
Ümit Ağbulut , Turkey
Najib Ben Aoun , Saudi Arabia
Surbhi Bhatia , Saudi Arabia
Daniele Bibbo , Italy
Vince D. Calhoun , USA
Francesco Camastra, Italy
Zhicheng Cao, China
Hubert Cecotti , USA
Jyotir Moy Chatterjee , Nepal
Rupesh Chikara, USA
Marta Cimitile, Italy
Silvia Conforto , Italy
Paolo Crippa , Italy
Christian W. Dawson, United Kingdom
Carmen De Maio , Italy
Thomas DeMarse , USA
Maria Jose Del Jesus, Spain
Arnaud Delorme , France
Anastasios D. Doulamis, Greece
António Dourado , Portugal
Sheng Du , China
Said El Kafhali , Morocco
Mohammad Reza Feizi Derakhshi , Iran
Quanxi Feng, China
Zhong-kai Feng, China
Steven L. Fernandes, USA
Agostino Forestiero , Italy
Piotr Franaszczuk , USA
Thippa Reddy Gadekallu , India
Paolo Gastaldo , Italy
Samanwoy Ghosh-Dastidar, USA

Manuel Graña , Spain
Alberto Guillén , Spain
Gaurav Gupta, India
Rodolfo E. Haber , Spain
Usman Habib , Pakistan
Anandakumar Haldorai , India
José Alfredo Hernández-Pérez , Mexico
Luis Javier Herrera , Spain
Alexander Hošovský , Slovakia
Etienne Hugues, USA
Nadeem Iqbal , Pakistan
Sajad Jafari, Iran
Abdul Rehman Javed , Pakistan
Jing Jin , China
Li Jin, United Kingdom
Kanak Kalita, India
Ryotaro Kamimura , Japan
Pasi A. Karjalainen , Finland
Anitha Karthikeyan, Saint Vincent and the Grenadines
Elpida Keravnou , Cyprus
Asif Irshad Khan , Saudi Arabia
Muhammad Adnan Khan , Republic of Korea
Abbas Khosravi, Australia
Tai-hoon Kim, Republic of Korea
Li-Wei Ko , Taiwan
Raşit Köker , Turkey
Deepika Koundal , India
Sunil Kumar , India
Fabio La Foresta, Italy
Kuruva Lakshmana , India
Maciej Lawrynczuk , Poland
Jianli Liu , China
Giosuè Lo Bosco , Italy
Andrea Loddo , Italy
Kezhi Mao, Singapore
Paolo Massobrio , Italy
Gerard McKee, Nigeria
Mohit Mittal , France
Paulo Moura Oliveira , Portugal
Debajyoti Mukhopadhyay , India
Xin Ning , China
Nasimul Noman , Australia
Fivos Panetsos , Spain

Evgeniya Pankratova , Russia
Rocío Pérez de Prado , Spain
Francesco Pistolesi , Italy
Alessandro Sebastian Podda , Italy
David M Powers, Australia
Radu-Emil Precup, Romania
Lorenzo Putzu, Italy
S P Raja, India
Dr.Anand Singh Rajawat , India
Simone Ranaldi , Italy
Upaka Rathnayake, Sri Lanka
Navid Razmjooy, Iran
Carlo Ricciardi, Italy
Jatinderkumar R. Saini , India
Sandhya Samarasinghe , New Zealand
Friedhelm Schwenker, Germany
Mijanur Rahaman Seikh, India
Tapan Senapati , China
Mohammed Shuaib , Malaysia
Kamran Siddique , USA
Gaurav Singal, India
Akansha Singh , India
Chiranjibi Sitaula , Australia
Neelakandan Subramani, India
Le Sun, China
Rawia Tahrir , Iraq
Binhua Tang , China
Carlos M. Travieso-González , Spain
Vinh Truong Hoang , Vietnam
Fath U Min Ullah , Republic of Korea
Pablo Varona , Spain
Roberto A. Vazquez , Mexico
Mario Versaci, Italy
Gennaro Vessio , Italy
Ivan Volosyak , Germany
Leyi Wei , China
Jianghui Wen, China
Lingwei Xu , China
Cornelio Yáñez-Márquez, Mexico
Zaher Mundher Yaseen, Iraq
Yugen Yi , China
Qiangqiang Yuan , China
Miaolei Zhou , China
Michal Zochowski, USA
Rodolfo Zunino, Italy

Contents

Retracted: Automatic Target Recognition of SAR Images Using Collaborative Representation

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9765863, Volume 2023 (2023)

Retracted: CT-ML: Diagnosis of Breast Cancer Based on Ultrasound Images and Time-Dependent Feature Extraction Methods Using Contourlet Transformation and Machine Learning

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9814037, Volume 2023 (2023)

Retracted: Visual Evaluation of Urban Streetscape Design Supported by Multisource Data and Deep Learning

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9853805, Volume 2023 (2023)

Retracted: Intelligent Recognition Model of Business English Translation Based on Improved GLR Algorithm

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9843471, Volume 2023 (2023)

Retracted: Rice Drought Damage Assessment Using AMSR-E Data Inversion of Surface Temperature

Computational Intelligence and Neuroscience




Retraction (1 page), Article ID 9835946, Volume 2023 (2023)

Retracted: Realization of Super-Large-Diameter Slurry Shield Passing through Settlement-Sensitive Area Based on Unreinforced Disturbance Control Technology

Computational Intelligence and Neuroscience




Retraction (1 page), Article ID 9786867, Volume 2023 (2023)

Computing Topological Invariants of Deep Neural Networks

Xiujun Zhang , Nazeran Idrees , Salma Kanwal , Muhammad Jawwad Saif, and Fatima Saeed



Research Article (11 pages), Article ID 9051908, Volume 2022 (2022)

Computing the Entropy Measures for the Line Graphs of Some Chemical Networks

Muhammad Farhan Hanif , Hasan Mahmood, Shazia Manzoor , and Fikre Bogale Petros 


Research Article (18 pages), Article ID 2006574, Volume 2022 (2022)

Generalized β -Hyers-Ulam-Rassias Stability of Impulsive Difference Equations

Yahya Almalki , Gul Rahmat, Atta Ullah, Fatima Shehryar, Muhammad Numan, and Muhammad Usman Ali 






Research Article (12 pages), Article ID 9462424, Volume 2022 (2022)

Key Information Extraction Algorithm of Different Types of Digital Archives for Cultural Operation and Management


Xiulun Ma 

Research Article (10 pages), Article ID 3459605, Volume 2022 (2022)


A Framework for Interactive Medical Image Segmentation Using Optimized Swarm Intelligence with Convolutional Neural Networks

Chetna Kaushal , Md Khairul Islam , Sara A. Althubiti , Fayadh Alenezi , and Romany F. Mansour 
Research Article (21 pages), Article ID 7935346, Volume 2022 (2022)

Intangible Cultural Heritage Reproduction and Revitalization: Value Feedback, Practice, and Exploration Based on the IPA Model

Lingling Xiao 
Research Article (13 pages), Article ID 8411999, Volume 2022 (2022)




Study on the Innovative Development of Digital Media Art in the Context of Artificial Intelligence

Chaomiao Chen 
Research Article (12 pages), Article ID 1004204, Volume 2022 (2022)


Optimization Mold and Algorithm of Risk Control for Power Grid Corporations Based on Collaborative Filtering Technology

Longxing Chen and Ping Han 
Research Article (9 pages), Article ID 3319311, Volume 2022 (2022)


Study of HCP (Hexagonal Close-Packed) Crystal Structure Lattice through Topological Descriptors

Guoping Zhang, Saadia Saeed, Adnan Aslam , Salma Kanwal , Nazeran Idrees , and Tahira Sumbal Shaikh
Research Article (13 pages), Article ID 6069666, Volume 2022 (2022)

AIOps Architecture in Data Center Site Infrastructure Monitoring

Wei Dong 
Research Article (12 pages), Article ID 1988990, Volume 2022 (2022)


Deep-Learning-Based Financial Message Sentiment Classification in Business Management

Chen Shao  and Xiaochen Chen
Research Article (9 pages), Article ID 3888675, Volume 2022 (2022)


Application of Data Mining Technology in Enterprise Green Innovation Model Construction and Path Analysis

Binfeng He 
Research Article (9 pages), Article ID 7194171, Volume 2022 (2022)

Deep Learning-Based Football Player Detection in Videos

Tianyi Wang  and Tongyan Li
Research Article (8 pages), Article ID 3540642, Volume 2022 (2022)

Recommendation of Business Models for Agriculture-Related Platforms Based on Deep Learning

Yufei Zhou and Sha Hua 
Research Article (5 pages), Article ID 7330078, Volume 2022 (2022)


Contents

Music Score Recognition Method Based on Deep Learning

Qin Lin 


Research Article (12 pages), Article ID 3022767, Volume 2022 (2022)

Smart Home Product Layout Design Method Based on Real-Number Coding Genetic Algorithm

Nan Jiang 




Research Article (10 pages), Article ID 1523330, Volume 2022 (2022)

Visual Network Analysis Based on Stereo Vision and Feature Matching Algorithm

Lili Shi 

Research Article (7 pages), Article ID 2910531, Volume 2022 (2022)

On Trees with Greatest F -Invariant Using Edge Swapping Operations

Wenhu Wang, Adnan Aslam , Muhammad Ahsan Binyamin , Salma Kanwal , and Iqra Irshad

Research Article (8 pages), Article ID 8291974, Volume 2022 (2022)

Development of Water Culture Tourism of Mountain Ethnic Culture Based on 3D Image Technology

Chuangle Guo  and Xinlu Shi



Research Article (9 pages), Article ID 5465488, Volume 2022 (2022)

Performance Evaluation Model of Agricultural Enterprise Technology Innovation Based on GA-BP Neural Network

Jian Kang and Minjuan Zhao 

Research Article (8 pages), Article ID 7110502, Volume 2022 (2022)

Robust Parameter Design for Cyclone System Based on Dual-Response Surface Method and Multiobjective Genetic Algorithm

Fusheng Luo , Xianhui Yin , and Zhanwen Niu


Research Article (12 pages), Article ID 5884868, Volume 2022 (2022)

Head CT Image Segmentation and Three-Dimensional Reconstruction Technology Based on Human Anatomy

Zhenyu Wu , Lin Wang , Yifei Li , Shuhui Dai , and Dongliang Zhang 


Research Article (10 pages), Article ID 7091476, Volume 2022 (2022)

Speckle Noise Removal Model Based on Diffusion Equation and Convolutional Neural Network

Siwei Nao and Yan Wang 


Research Article (11 pages), Article ID 5344263, Volume 2022 (2022)

Mask R-CNN-Oriented Pottery Display and Identification System

Chuantao Wei 


Research Article (9 pages), Article ID 6288201, Volume 2022 (2022)

Protection and Planning of Historic Districts Based on Internet of Things Perception

Qing Zhao  and YunYing Ren


Research Article (12 pages), Article ID 1325381, Volume 2022 (2022)

MOOC Teaching Model of Basic Education Based on Fuzzy Decision Tree Algorithm

Zhang Yuanyuan 







Research Article (7 pages), Article ID 3175028, Volume 2022 (2022)

Generative Adversarial Network for Musical Notation Recognition during Music Teaching

Na Li 


Research Article (9 pages), Article ID 8724688, Volume 2022 (2022)

Prediction of Diabetes through Retinal Images Using Deep Neural Network

Mahmoud Ragab , Abdullah S. AL-Malaise AL-Ghamdi , Bahjat Fakieh , Hani Choudhry , Romany F. Mansour , and Deepika Koundal 


Research Article (6 pages), Article ID 7887908, Volume 2022 (2022)

The Collection and Recognition Method of Music and Dance Movement Based on Intelligent Sensor

Jun Geng 


Research Article (9 pages), Article ID 2654892, Volume 2022 (2022)

Exploring the Integration Model of Industry Chain Information System Based on Energy Internet and Its Key Technologies

Anxun Qi 


Research Article (10 pages), Article ID 8752048, Volume 2022 (2022)

Comprehensive Evaluation Method of Digital Economy Development Level Based on Complex Network Model

Xingjun Shi 


Research Article (8 pages), Article ID 4999178, Volume 2022 (2022)

Construction of Remote Sensing Model of Fresh Corn Biomass Based on Neural Network

Jianjian Chen, Hui Zhang, Yunlong Bian, Xiangnan Li, and Guihua Lv 

Research Article (8 pages), Article ID 2844563, Volume 2022 (2022)

The Research of Chinese Martial Arts Cross-Media Communication System Based on Deep Neural Network

Yue Su , Jing Tian, and Xin Zan


Research Article (9 pages), Article ID 2835992, Volume 2022 (2022)

[Retracted] CT-ML: Diagnosis of Breast Cancer Based on Ultrasound Images and Time-Dependent Feature Extraction Methods Using Contourlet Transformation and Machine Learning

Behnam Hajipour Khire Masjidi, Soufia Bahmani , Fatemeh Sharifi, Mohammad Peivandi, Mohammad Khosravani, and Adil Hussein Mohammed

Research Article (15 pages), Article ID 1493847, Volume 2022 (2022)


[Retracted] Automatic Target Recognition of SAR Images Using Collaborative Representation

Jinge Hu 

Research Article (7 pages), Article ID 3100028, Volume 2022 (2022)

Contents

Construction of Enterprise Financial Early Warning Model Based on Logistic Regression and BP Neural Network

Jincheng Lyu 


Research Article (7 pages), Article ID 2614226, Volume 2022 (2022)

Research on Intelligent Scheduling Scheme of Aerobics Competition for Multi-Intelligent Decision-Making

Hongjing Guan , Zhi Tan, and Xingrui Zhao


Research Article (12 pages), Article ID 3407494, Volume 2022 (2022)

Voice Recognition and Evaluation of Vocal Music Based on Neural Network

Xiaochen Wang and Tao Wang 


Research Article (9 pages), Article ID 3466987, Volume 2022 (2022)

[Retracted] Rice Drought Damage Assessment Using AMSR-E Data Inversion of Surface Temperature

Ying Wang, Jianping Zhang , and Guochun Li

Research Article (8 pages), Article ID 4117481, Volume 2022 (2022)

Intelligent Classification Method of Archive Data Based on Multigranular Semantics

Xiaobo Jiang 

Research Article (9 pages), Article ID 7559523, Volume 2022 (2022)

Analysis and Prediction of Subway Tunnel Surface Subsidence Based on Internet of Things Monitoring and BP Neural Network

Baitian Wang, Jing Zhang, Longhao Zhang, Shi Yan , Qiangqiang Ma , Wentao Li, and Maopeng Jiao

Research Article (10 pages), Article ID 9447897, Volume 2022 (2022)

Automatic Recognition Method of Machine English Translation Errors Based on Multisignal Feature Fusion

Ruisi Zhang  and Haibo Huang

Research Article (9 pages), Article ID 2987227, Volume 2022 (2022)

Optimization of Reservoir Flood Control Operation Based on Multialgorithm Deep Learning

Bowen Xue , Yan Xie, Yanhui Liu , Along Li, Daguang Zhao, and Haipeng Li

Research Article (10 pages), Article ID 4123421, Volume 2022 (2022)

Prediction of Response to Radiotherapy by Characterizing the Transcriptomic Features in Clinical Tumor Samples across 15 Cancer Types

Yu Xu, Chao Tang, Yan Wu, Ling Luo, Ying Wang, Yongzhong Wu , and Xiaolong Shi 


Research Article (16 pages), Article ID 5443709, Volume 2022 (2022)

Detection Anomaly in Video Based on Deep Support Vector Data Description

Bokun Wang, Caiqian Yang , and Yaojing Chen







Research Article (6 pages), Article ID 5362093, Volume 2022 (2022)

Improved Deep Neural Network for Cross-Media Visual Communication

Yubo Miao 

Research Article (9 pages), Article ID 1556352, Volume 2022 (2022)

Hyperspectral Image Classification: Potentials, Challenges, and Future Directions

Debaleena Datta , Pradeep Kumar Mallick , Akash Kumar Bhoi , Muhammad Fazal Ijaz , Jana Shafi , and Jaeyoung Choi 


Review Article (36 pages), Article ID 3854635, Volume 2022 (2022)

English Teaching Quality Monitoring and Multidimensional Analysis Based on the Internet of Things and Deep Learning Model

Juan Song 






Research Article (10 pages), Article ID 9667864, Volume 2022 (2022)

Automatic Gray Image Coloring Method Based on Convolutional Network

Jiayi Fan , Wentao Xie, and Tiantian Ge


Research Article (9 pages), Article ID 5273698, Volume 2022 (2022)

A Multitask Deep Learning Framework for DNER

Ran Jin , Tengda Hou , Tongrui Yu , Min Luo , and Haoliang Hu 

Research Article (10 pages), Article ID 3321296, Volume 2022 (2022)

Prediction of Purchase Volume of Cross-Border e-Commerce Platform Based on BP Neural Network

Xiang Zhang 



Research Article (9 pages), Article ID 3821642, Volume 2022 (2022)

[Retracted] Intelligent Recognition Model of Business English Translation Based on Improved GLR Algorithm

Laiying Deng, Xinde Hu, and Fen Liu 


Research Article (8 pages), Article ID 4105942, Volume 2022 (2022)

A Novel Encoder-Decoder Model for Multivariate Time Series Forecasting

Huihui Zhang , Shicheng Li , Yu Chen , Jiangyan Dai , and Yugen Yi 




Research Article (17 pages), Article ID 5596676, Volume 2022 (2022)

Algorithm of CAD Surface Generation for Complex Pipe Model in Industry 4.0 Background

Xiaolei Cheng 

Research Article (12 pages), Article ID 7062052, Volume 2022 (2022)





Discrete Dynamic Modeling of Learner Behavior Analysis in Physical Education Teaching

Jia Shi , Jun Sun , and Zhonghua Zheng 

Research Article (7 pages), Article ID 4620599, Volume 2022 (2022)

Contents

SCU-Net: Semantic Segmentation Network for Learning Channel Information on Remote Sensing Images

Wei Wang , Yuxi Kang , Guanqun Liu , and Xin Wang 
Research Article (11 pages), Article ID 8469415, Volume 2022 (2022)

Research on Urban National Sports Fitness Demand Prediction Method Based on Ant Colony Algorithm

Wei Yue  and Peng Dai 
Research Article (9 pages), Article ID 5872643, Volume 2022 (2022)

3D Virtual Modeling Realizations of Building Construction Scenes via Deep Learning Technique

Weihong Li 
Research Article (11 pages), Article ID 6286420, Volume 2022 (2022)

Dynamic Analysis of Deep Water Highway Tunnel under Ocean Current

Li Fang , Hong Li , and Bin Li
Research Article (18 pages), Article ID 9551792, Volume 2022 (2022)

Spatial-Temporal Graph Convolutional Framework for Yoga Action Recognition and Grading

Shu Wang 
Research Article (9 pages), Article ID 7500525, Volume 2022 (2022)


China's Economic Forecast Based on Machine Learning and Quantitative Easing

Chang Qiu 
Research Article (8 pages), Article ID 2404174, Volume 2022 (2022)

An Intelligent Classification Method of Multisource Enterprise Financial Data Based on SAS Model

Xiuyan Xu 
Research Article (9 pages), Article ID 8255091, Volume 2022 (2022)

5G Network Slicing: Methods to Support Blockchain and Reinforcement Learning

Juan Hu  and Jianwei Wu
Research Article (10 pages), Article ID 1164273, Volume 2022 (2022)


An Empirical Analysis of the Influence of Volleyball Elective Course on Students' Physical Health Based on Digital Image

Shiwei Wang 
Research Article (15 pages), Article ID 9229912, Volume 2022 (2022)

A Data Augmentation Method for Prohibited Item X-Ray Pseudocolor Images in X-Ray Security Inspection Based on Wasserstein Generative Adversarial Network and Spatial-and-Channel Attention Block


Dongming Liu , Jianchang Liu , Peixin Yuan , and Feng Yu 
Research Article (14 pages), Article ID 8172466, Volume 2022 (2022)

Prediction Model of Stress Intensity Factor of Circumferential Through Crack in Elbow Based on Neural Network

Xiaohong Li , Xianghui Li, and Bin Chen



Research Article (6 pages), Article ID 8395505, Volume 2022 (2022)

Design of Fault Prediction System for Electromechanical Sensor Equipment Based on Deep Learning

Yongtao Ding , Hua Wu, and Kaixiang Zhou


Research Article (10 pages), Article ID 3057167, Volume 2022 (2022)

Dynamic Evolution Analysis of Desertification Images Based on BP Neural Network

Guanyao Lu , Dan Xu, and Yue Meng 


Research Article (11 pages), Article ID 5645535, Volume 2022 (2022)

Human Resource Planning and Configuration Based on Machine Learning

Shuai Yuan, Qian Qi , Enliang Dai, and Yongfeng Liang

Research Article (6 pages), Article ID 3605722, Volume 2022 (2022)

English Text Readability Measurement Based on Convolutional Neural Network: A Hybrid Network Model

Lihua Jian , Huiqun Xiang, and Guobin Le


Research Article (9 pages), Article ID 6984586, Volume 2022 (2022)

Coal Mine Safety Evaluation Based on Machine Learning: A BP Neural Network Model

Guangxing Bai  and Tianlong Xu


Research Article (9 pages), Article ID 5233845, Volume 2022 (2022)

Aided Recognition and Training of Music Features Based on the Internet of Things and Artificial Intelligence

Xidan Zhang 


Research Article (11 pages), Article ID 3733818, Volume 2022 (2022)

Collaborative Filtering Recommendation of Music MOOC Resources Based on Spark Architecture

Lifu Wang 


Research Article (8 pages), Article ID 2117081, Volume 2022 (2022)

Research on Ecoenvironmental Quality Evaluation System Based on Big Data Analysis

Pingheng Li 

Research Article (14 pages), Article ID 5191223, Volume 2022 (2022)


Deep Learning-Based Fake Information Detection and Influence Evaluation

Ning Xiang 

Research Article (8 pages), Article ID 8514430, Volume 2022 (2022)


Contents

Research on the Construction of College Football Classroom Practice Teaching System Model Based on Big Data Analysis

Bo Zhang and Wei Ren 

Research Article (11 pages), Article ID 5018033, Volume 2022 (2022)

[Retracted] Visual Evaluation of Urban Streetscape Design Supported by Multisource Data and Deep Learning

Guanqing Feng, Guangtian Zou , and Pengjin Wang


Research Article (9 pages), Article ID 3287117, Volume 2022 (2022)

Intelligent Detection and Analysis of Polycyclic Aromatic Hydrocarbons Based on Surface-Enhanced Raman Scattering Spectroscopy

Qian Zhang , Bowen Chen, Fazli Wahid, Wanyun Feng, and Xuerou Chen


Research Article (7 pages), Article ID 8330702, Volume 2022 (2022)

Research on MOOC Teaching Mode in Higher Education Based on Deep Learning

Yuan Tian, Yingjie Sun, Lijing Zhang, and Wanqiang Qi 

Research Article (10 pages), Article ID 8031602, Volume 2022 (2022)

The Design of Adolescents' Physical Health Prediction System Based on Deep Reinforcement Learning

Hailiang Sun and Dan Yang 


Research Article (10 pages), Article ID 4946009, Volume 2022 (2022)

Collaborative Research on Mouth Shape and Lyrics in Singing Practice Based on Image Processing

Lujia Xu and Chen Chen 

Research Article (13 pages), Article ID 5138442, Volume 2022 (2022)

Construction of a Hybrid Teaching Model System Based on Promoting Deep Learning

Yingjie Sun, Yang Li, Yuan Tian, and Wanqiang Qi 

Research Article (12 pages), Article ID 4447530, Volume 2022 (2022)

Research on English Achievement Analysis Based on Improved CARMA Algorithm

Lin Hu 


Research Article (11 pages), Article ID 8687879, Volume 2022 (2022)

Application of Image Mosaic Technology in Tai Chi Animation Creation

Yajun Pang 

Research Article (9 pages), Article ID 4775189, Volume 2022 (2022)

Research on Teaching Practice of Blended Higher Education Based on Deep Learning Route

Yang Li, Lijing Zhang, Yuan Tian, and Wanqiang Qi 


Research Article (8 pages), Article ID 5906335, Volume 2022 (2022)

[Retracted] Realization of Super-Large-Diameter Slurry Shield Passing through Settlement-Sensitive Area Based on Unreinforced Disturbance Control Technology

Dongshuang Liu, Xinrong Liu , Zuliang Zhong , Yafeng Han, Fei Xiong, and Xiaohan Zhou 

Research Article (9 pages), Article ID 6299645, Volume 2022 (2022)

Risk Analysis of Textile Industry Foreign Investment Based on Deep Learning

Jingyi Liu and Jiaolong Li 

Research Article (6 pages), Article ID 3769670, Volume 2022 (2022)

Retraction

Retracted: Automatic Target Recognition of SAR Images Using Collaborative Representation

Computational Intelligence and Neuroscience

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

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

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

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

References

- [1] J. Hu, "Automatic Target Recognition of SAR Images Using Collaborative Representation," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3100028, 7 pages, 2022.

Retraction

Retracted: CT-ML: Diagnosis of Breast Cancer Based on Ultrasound Images and Time-Dependent Feature Extraction Methods Using Contourlet Transformation and Machine Learning

Computational Intelligence and Neuroscience

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

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

References

- [1] B. Hajipour Khire Masjidi, S. Bahmani, F. Sharifi, M. Peivandi, M. Khosravani, and A. Hussein Mohammed, "CT-ML: Diagnosis of Breast Cancer Based on Ultrasound Images and Time-Dependent Feature Extraction Methods Using Contourlet Transformation and Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1493847, 15 pages, 2022.

Retraction

Retracted: Visual Evaluation of Urban Streetscape Design Supported by Multisource Data and Deep Learning

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

References

- [1] G. Feng, G. Zou, and P. Wang, "Visual Evaluation of Urban Streetscape Design Supported by Multisource Data and Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3287117, 9 pages, 2022.

Retraction

Retracted: Intelligent Recognition Model of Business English Translation Based on Improved GLR Algorithm

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

References

- [1] L. Deng, X. Hu, and F. Liu, "Intelligent Recognition Model of Business English Translation Based on Improved GLR Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4105942, 8 pages, 2022.

Retraction

Retracted: Rice Drought Damage Assessment Using AMSR-E Data Inversion of Surface Temperature

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

References

- [1] Y. Wang, J. Zhang, and G. Li, "Rice Drought Damage Assessment Using AMSR-E Data Inversion of Surface Temperature," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4117481, 8 pages, 2022.

Retraction

Retracted: Realization of Super-Large-Diameter Slurry Shield Passing through Settlement-Sensitive Area Based on Unreinforced Disturbance Control Technology

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

References

- [1] D. Liu, X. Liu, Z. Zhong, Y. Han, F. Xiong, and X. Zhou, "Realization of Super-Large-Diameter Slurry Shield Passing through Settlement-Sensitive Area Based on Unreinforced Disturbance Control Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6299645, 9 pages, 2022.

Research Article

Computing Topological Invariants of Deep Neural Networks

Xiujun Zhang ¹, **Nazeran Idrees** ², **Salma Kanwal** ³, **Muhammad Jawwad Saif**,⁴
and **Fatima Saeed**²

¹*School of Computer Science, Chengdu University, Chengdu, China*

²*Department of Mathematics, Government College University Faisalabad, Faisalabad 38000, Pakistan*

³*Department of Mathematics, Lahore College for Women University, Lahore 54000, Pakistan*

⁴*Department of Applied Chemistry, Government College University Faisalabad, Faisalabad 38000, Pakistan*

Correspondence should be addressed to Nazeran Idrees; nazeranidrees@gcuf.edu.pk and Salma Kanwal; salma.kanwal055@gmail.com

Received 8 May 2022; Revised 27 July 2022; Accepted 12 September 2022; Published 7 October 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiujun 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.

A deep neural network has multiple layers to learn more complex patterns and is built to simulate the activity of the human brain. Currently, it provides the best solutions to many problems in image recognition, speech recognition, and natural language processing. The present study deals with the topological properties of deep neural networks. The topological index is a numeric quantity associated to the connectivity of the network and is correlated to the efficiency and accuracy of the output of the network. Different degree-related topological indices such as Zagreb index, Randić index, atom-bond connectivity index, geometric-arithmetic index, forgotten index, multiple Zagreb indices, and hyper-Zagreb index of deep neural network with a finite number of hidden layers are computed in this study.

1. Introduction

Neural networks are not only studied in artificial intelligence but also have got great applications in intrusion detection systems, image processing, localization, medicine, and chemical and environmental sciences [1–3]. Neural networks are used to model and learn complex and nonlinear relationships, which is very important in real life because many of the relationships of inputs and outputs are nonlinear and complex. Artificial neural networks are the backbone of robotics, defense technology, and neural chemistry. Neural networks are not only being widely used as a tool for predictive analysis but also trained successfully to model processes including crystallization, adsorption, distillation, gasification, dry reforming, and filtration in neural chemistry [4–8].

The topological index associates a unique number to a graph or network, which provides correlation with the physiochemical properties of the network. Degree-based topological index depends upon the connectivity of the

network. The first degree-based topological index, called the Randić index, was formulated by Milan Randić [9] while analyzing the boiling point of paraffin. Over the last three decades, hundreds of topological indices have been formulated by researchers, which are helpful in studying the different properties of chemical graphs like reactivity, stability, boiling point, enthalpy of formation, and Kovat's constant and inherits physical properties of materials such as stress, elasticity, strain, mechanical strength, and many others.

Bollobás and Erdős [10] introduced the general Randić index given by equation (1). The first and second Zagreb indices were introduced by Gutman and Trinajstić [11] in 1972, which appeared during the analysis of π -electron energy of atoms. The multiplicative version of these Zagreb indices (the first multiplicative Zagreb index and the second multiplicative Zagreb index) of a graph were formulated by Ghorbani and Azimi [12]. Shirdel et al. [13] introduced a new version of Zagreb indices named as the hyper-Zagreb index. The widely used atom-bond connectivity (ABC) index is

introduced by Estrada et al. [14]. Zhou and Trinajstić [15] gave the idea of the sum-connectivity index (SCI). The geometric-arithmetic index was introduced by Vukičević and Furtula [16]. Javaid et al. [17] investigated the degree-based topological indices for the probabilistic neural networks in 2017. Topological indices for multilayered probabilistic neural networks and recurrent neural networks have also been computed recently [18–21]. For more work-related to computation and bounds of topological indices, see [22–29].

Consider a graph G having a set of nodes V and a set of edges E . Degree of a node v , denoted by d_v , is the number of nodes connected to v via an edge. A degree-based topological indices of a graph G are defined as follows:

Randić index

$$\chi(G) = \sum_{uv \in E} \left(1/\sqrt{d_u d_v} \right). \quad (1)$$

General Randić index

$$R_\alpha(G) = \sum_{uv \in E} (d_u d_v)^\alpha. \quad (2)$$

First Zagreb index

$$M_1(G) = \sum_{uv \in E} [d_u + d_v]. \quad (3)$$

Second Zagreb index

$$M_2(G) = \sum_{uv \in E} [d_u d_v]. \quad (4)$$

First multiple Zagreb index

$$PM_1(G) = \prod_{uv \in E} [d_u \times d_v]. \quad (5)$$

Second multiple Zagreb index

$$PM_2(G) = \prod_{uv \in E} [d_u \times d_v]. \quad (6)$$

Hyper-Zagreb index

$$HM(C[m, n]) = \sum_{uv \in E} (d_u + d_v)^2. \quad (7)$$

Atom-bond connectivity index

$$ABC(G) = \sum_{uv \in E} \sqrt{(d_u + d_v - 2/d_u d_v)}. \quad (8)$$

Sum connectivity index

$$SCI(G) = \sum_{uv \in E} \left(1/\sqrt{d_u + d_v} \right). \quad (9)$$

Geometric-arithmetic index

$$GA(G) = \sum_{uv \in E} \left(2\sqrt{d_u d_v / (d_u + d_v)} \right). \quad (10)$$

2. Methodology

A deep neural network (DNN) can be represented by a graph $Z = (V, E)$, where V denotes the nodes of the network

and E denotes the set of edges between the nodes. We consider a DNN with an input layer having M nodes, r hidden layers each layer having $N_i, i = 1, 2, \dots, r$ number of nodes such that the first layer has N_1 nodes, the second layer has N_2 nodes, and similarly, the r -th layer has N_r nodes, which can also be expressed as $DNN(N_1 N_2 \dots N_r)$. The output layer of DNN has N nodes. Each node of every layer is connected to all nodes of the next layer. For instance, Figure 1 shows a DNN with an input layer having four nodes, an output layer with three nodes, and five hidden layers.

We first partition the edges of the graph of DNN according to the degree of end vertices of the graph. We analyze the structure of the graph by considering the connectivity of vertices of each layer to the next layer. In DNN, each node of every layer is connected to all nodes of the next layer. This fact is employed to count the degree of each vertex. Consider a deep neural network $DNN(N_1 N_2 \dots N_r)$. Each node in the input layer has a degree N_1 because every input node is connected to each node of a first hidden layer having N_1 nodes. In the first hidden layer, all node (N_1) has the same degree, i.e., $M + N_2$. Nodes of the second layer have degree $N_1 + N_3$. Similarly, the nodes of i -th hidden layer have degree $N_{i-1} + N_{i+1}$. The nodes of the output layer have degree N_r .

We will compute topological indices using the edge partition method. We will classify the edges on basis of degrees of end-nodes of the edges. The number of edges connecting the input layer to the first hidden layer is $N_1 M$, whose end-nodes have degrees N_1 and $M + N_2$. The edges connecting i -th hidden layer to $i + 1$ -st layer have end-nodes having degrees $N_{i-1} + N_{i+1}$ and $N_i + N_{i+2}$ and the number of such edges is $N_i N_{i+1}$. Similarly, the $N_r N$ edges connecting the last hidden layer to the output layer have degrees $N_{r-1} + N$ and N_r of end-nodes. These findings are summarized in Table 1 below, which will be further helpful in computing the topological indices.

3. Results and Discussions

In this section, we have derived the expressions to compute the topological indices of the deep neural network. These results are related to the connectivity of nodes of DNN.

Theorem 1. Let $Z \cong DNN(N_1 N_2 \dots N_r)$ be a deep neural network. Then the Randić index ($R_{(1/2)}(Z)$) and general Randić index ($R_\alpha(Z)$) of DNN are given as

$$\begin{aligned} (i) \quad R_{(1/2)}(Z) &= (MN_1)\sqrt{(N_1)(M + N_2)} + (N_1 N_2) \\ &\quad \sqrt{(M + N_2)(N_1 + N_3)} + \left(\sum_{i=2}^{r-2} N_i N_{i+1} ((N_{i-1} + N_{i+1})(N_i + N_{i+2}))^{(1/2)} \right) \\ &\quad + (N_{r-1} N_r) \sqrt{(N_{r-2} + N_r)(N_{r-1} + N)} + \\ &\quad (N_r N) \sqrt{(N_{r-1} + N)(N_r)}, \\ (ii) \quad R_\alpha(Z) &= (MN_1)((N_1)(M + N_2))^\alpha + (N_1 N_2) \\ &\quad ((M + N_2)(N_1 + N_3))^\alpha + \left(\sum_{i=2}^{r-2} N_i N_{i+1} ((N_{i-1} + N_{i+1})(N_i + N_{i+2}))^\alpha \right) \\ &\quad + (N_{r-1} N_r) ((N_{r-2} + N_r)(N_{r-1} + N))^\alpha + (N_r N) ((N_{r-1} + N)(N_r))^\alpha. \end{aligned}$$

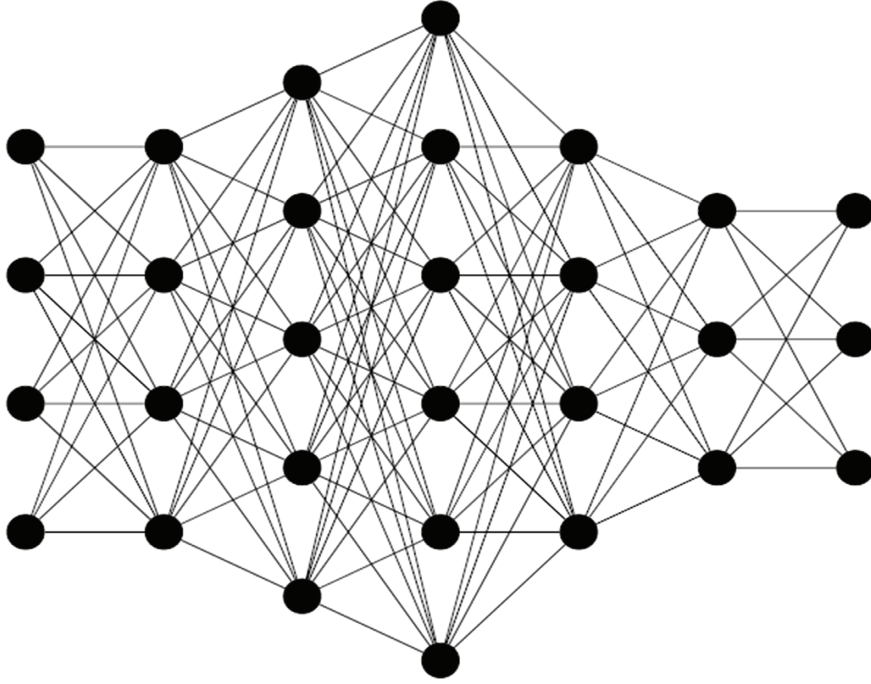


FIGURE 1: A deep neural network with five hidden layers of DNN (4,4,5,6,4,3,3).

 TABLE 1: The edge partition of $DNN(N_1N_2 \dots N_r)$ based on degrees of end nodes.

$(d_u, d_v), uv \in E(Z)$	Number of edges of the form (d_u, d_v)
$(N_1, M + N_2)$	MN_1
$(M + N_2, N_1 + N_3)$	N_1N_2
$(N_1 + N_3, N_2 + N_4)$	N_2N_3
\vdots	\vdots
$(N_{r-3} + N_{r-1}, N_{r-2} + N_r)$	$N_{r-2}N_{r-1}$
$(N_{r-2} + N_r, N_{r-1} + N)$	$N_{r-1}N_r$
$(N_{r-1} + N, N_r)$	N_rN

Proof. We calculated the degrees of end nodes of every edge for $DNN(N_1N_2 \dots N_r)$. By using the definitions and values from Table 1, we get the following results:

$$\begin{aligned}
 \text{(i) } R_{(1/2)}(Z) &= \sum_{uv \in E} \sqrt{d_u d_v} = \sum_{uv \in E(N_1, M+N_2)} \sqrt{d_u d_v} + \\
 &\sum_{uv \in E(M+N_2, N_1+N_3)} \sqrt{d_u d_v} + \sum_{uv \in E(N_1+N_3, N_2+N_4)} \sqrt{d_u d_v} \dots \\
 &\sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} \sqrt{d_u d_v} + \\
 &\sum_{uv \in E((N_{r-2}+N_r, N_{r-1}+N))} \sqrt{d_u d_v} + \sum_{uv \in E(N_{r-1}+N, N_r)} \sqrt{d_u d_v}
 \end{aligned}$$

Substituting values from Table 1, we get

$$\begin{aligned}
 R_{(1/2)}(Z) &= (MN_1)\sqrt{(N_1)(M + N_2)} \\
 &+ (N_1N_2)\sqrt{(M + N_2)(N_1+N_3) + (N_2N_3)} \\
 &\sqrt{(N_1 + N_3)(N_2 + N_4)} + \dots \\
 &+ (N_{r-2}N_{r-1})\sqrt{(N_{r-3}+N_{r-1})(N_{r-2}+N_r)} \\
 &+ (N_{r-1}N_r)\sqrt{(N_{r-2} + N_r)(N_{r-1} + N)} \\
 &+ (N_rN)\sqrt{(N_{r-1} + N)(N_r)}.
 \end{aligned} \tag{11}$$

This can be expressed as follows:

$$\begin{aligned}
 R_{(1/2)}(Z) &= (MN_1)\sqrt{(N_1)(M + N_2)} \\
 &+ (N_1N_2)\sqrt{(M + N_2)(N_1+N_3)} \\
 &+ \sum_{i=2}^{r-2} N_i N_{i+1} ((N_{i-1} + N_{i+1})(N_i + N_{i+2}))^{(1/2)} \\
 &+ (N_{r-1}N_r)\sqrt{(N_{r-2} + N_r)(N_{r-1} + N)} \\
 &+ (N_rN)\sqrt{(N_{r-1} + N)(N_r)}.
 \end{aligned} \tag{12}$$

$$(ii) R_\alpha(Z) = \sum_{uv \in E(N_1, M+N_2)} (d_u d_v)^\alpha + \sum_{uv \in E(M+N_2, N_1+N_3)} (d_u d_v)^\alpha + \sum_{uv \in E(N_1+N_3, N_2+N_4)} (d_u d_v)^\alpha + \dots + \sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} (d_u d_v)^\alpha + \sum_{uv \in E((N_{r-2}+N_r, N_{r-1}+N))} (d_u d_v)^\alpha + \sum_{uv \in E(N_{r-1}+N, N_r)} (d_u d_v)^\alpha$$

Using Table 1, we get

$$R_\alpha(Z) = (MN_1)((N_1)(M+N_2))^\alpha + (N_1N_2)((M+N_2)(N_1+N_3))^\alpha + (N_2N_3)((N_1+N_3)(N_2+N_4))^\alpha + \dots + (N_{r-2}N_{r-1})((N_{r-3}+N_{r-1})(N_{r-2}+N_r))^\alpha + (N_{r-1}N_r)((N_{r-2}+N_r)(N_{r-1}+N))^\alpha + (N_rN)((N_{r-1}+N)(N_r))^\alpha. \quad (13)$$

This can be further summarized as

$$R_\alpha(Z) = (MN_1)((N_1)(M+N_2))^\alpha + (N_1N_2)((M+N_2)(N_1+N_3))^\alpha + \sum_{i=2}^{r-2} N_i N_{i+1} ((N_{i-1}+N_{i+1})(N_i+N_{i+2}))^\alpha + (N_{r-1}N_r)((N_{r-2}+N_r)(N_{r-1}+N))^\alpha + (N_rN)((N_{r-1}+N)(N_r))^\alpha. \quad (14)$$

□

Theorem 2. Let $Z \cong DNN(N_1N_2 \dots N_r)$ be a deep neural network. Then, first Zagreb index ($M_1(Z)$), second Zagreb index ($M_2(Z)$), first multiplicative Zagreb ($PM_1(Z)$) index, and second multiplicative Zagreb index ($PM_2(Z)$) of DNN are given as follows:

- (i) $M_1(Z) = (MN_1)(N_1 + M + N_2) + (N_1N_2)(M + N_2 + N_1 + N_3) + \sum_{i=2}^{r-2} N_i N_{i+1} (N_{i-1} + N_{i+1} + N_i + N_{i+2}) + (N_{r-1}N_r)(N_{r-2} + N_r + N_{r-1} + N) + (N_rN)(N_{r-1} + N + N_r)$
- (ii) $M_2(Z) = (MN_1)[(N_1)(M+N_2)] + (N_1N_2)[(M+N_2)(N_1+N_3)] + \sum_{i=2}^{r-2} N_i N_{i+1} [(N_{i-1}+N_{i+1})(N_i+N_{i+2})] + (N_{r-1}N_r)[(N_{r-2}+N_r)(N_{r-1}+N)] + (N_rN)[(N_{r-1}+N)(N_r)]$
- (iii) $PM_1(Z) = (N_1 + M + N_2)^{MN_1} \times (M + N_2 + N_1 + N_3)^{N_1N_2} \times \prod_{i=2}^{r-2} (N_{i-1} + N_{i+1} + N_i + N_{i+2})^{N_i N_{i+1}} \times (N_{r-2} + N_r + N_{r-1} + N)^{N_{r-1}N_r} \times (N_{r-1} + N + N_r)^{N_rN}$
- (iv) $PM_2(Z) = [(N_1)(M+N_2)]^{MN_1} \times [(M+N_2)(N_1+N_3)]^{N_1N_2} \times \prod_{i=2}^{r-2} [(N_{i-1}+N_{i+1})(N_i+N_{i+2})]^{N_i N_{i+1}} (N_i + N_{i+2})^{N_i N_{i+1}} \times [(N_{r-2}+N_r)(N_{r-1}+N)]^{N_{r-1}N_r} \times [(N_{r-1}+N)(N_r)]^{N_rN}$

$$N_{i+2})^{N_i N_{i+1}} \times [(N_{r-2} + N_r)(N_{r-1} + N)]^{N_{r-1}N_r} \times [(N_{r-1} + N)(N_r)]^{N_rN}$$

Proof. To compute the topological indices of DNN, we use the edge partition method. In Table 1, we have calculated the degrees of end-nodes of each edge for $DNN(N_1N_2 \dots N_r)$. Now, by using the definitions and values from Table 1, we have the following results:

$$(i) M_1(Z) = \sum_{uv \in E} [d_u + d_v] = \sum_{uv \in E(N_1, M+N_2)} [d_u + d_v] + \sum_{uv \in E(M+N_2, N_1+N_3)} [d_u + d_v] + \sum_{uv \in E(N_1+N_3, N_2+N_4)} [d_u + d_v] + \dots + \sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} [d_u + d_v] + \sum_{uv \in E((N_{r-2}+N_r, N_{r-1}+N))} [d_u + d_v] + \sum_{uv \in E(N_{r-1}+N, N_r)} [d_u + d_v]$$

Substituting values from Table 1, we get

$$M_1(Z) = (MN_1)(N_1 + M + N_2) + (N_1N_2)(M + N_2 + N_1 + N_3) + (N_2N_3)(N_1 + N_3 + N_2 + N_4) + \dots + (N_{r-2}N_{r-1})(N_{r-3} + N_{r-1} + N_{r-2} + N_r) + (N_{r-1}N_r)(N_{r-2} + N_r + N_{r-1} + N) + (N_rN)(N_{r-1} + N + N_r). \quad (15)$$

It can be expressed as follows:

$$M_1(Z) = (MN_1)(N_1 + M + N_2) + (N_1N_2)(M + N_2 + N_1 + N_3) + \sum_{i=2}^{r-2} N_i N_{i+1} (N_{i-1} + N_{i+1} + N_i + N_{i+2}) + (N_{r-1}N_r)(N_{r-2} + N_r + N_{r-1} + N) + (N_rN)(N_{r-1} + N + N_r). \quad (16)$$

$$(ii) M_2(Z) = \sum_{uv \in E} [d_u d_v] = \sum_{uv \in E(N_1, M+N_2)} [d_u d_v] + \sum_{uv \in E(M+N_2, N_1+N_3)} [d_u d_v] + \sum_{uv \in E(N_1+N_3, N_2+N_4)} [d_u d_v] + \dots + \sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} [d_u d_v] + \sum_{uv \in E((N_{r-2}+N_r, N_{r-1}+N))} [d_u d_v] + \sum_{uv \in E(N_{r-1}+N, N_r)} [d_u d_v]$$

Substituting values from Table 1, we have

$$\begin{aligned}
M_2(Z) &= (MN_1)[(N_1)(M+N_2)] \\
&+ (N_1N_2)[(M+N_2)(N_1+N_3)] \\
&+ (N_2N_3)[(N_1+N_3)(N_2+N_4)] + \dots \\
&+ (N_{r-2}N_{r-1})[(N_{r-3}+N_{r-1})(N_{r-2}+N_r)] \\
&+ (N_{r-1}N_r)[(N_{r-2}+N_r)(N_{r-1}+N)] \\
&+ (N_rN)[(N_{r-1}+N)(N_r)]. \tag{17}
\end{aligned}$$

It can be expressed as follows:

$$\begin{aligned}
M_2(Z) &= (MN_1)[(N_1)(M+N_2)] \\
&+ (N_1N_2)[(M+N_2)(N_1+N_3)] \\
&+ \sum_{i=2}^{r-2} N_iN_{i+1}(N_{i-1}+N_{i+1})(N_i+N_{i+2}) \tag{18} \\
&+ (N_{r-1}N_r)[(N_{r-2}+N_r)(N_{r-1}+N)] \\
&+ (N_rN)[(N_{r-1}+N)(N_r)].
\end{aligned}$$

$$\begin{aligned}
\text{(iii) } PM_1(Z) &= \prod_{uv \in E} [d_u + d_v] = \prod_{uv \in E(N_1, M+N_2)} [d_u + d_v] \\
&+ \prod_{uv \in E(M+N_2, N_1+N_3)} [d_u + d_v] + \prod_{uv \in E(N_1+N_3, N_2+N_4)} \\
&[d_u + d_v] \dots + \prod_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} [d_u + d_v] + \\
&\prod_{uv \in E(N_{r-2}+N_r, N_{r-1}+N)} [d_u + d_v] + \prod_{uv \in E(N_{r-1}+N, N_r)} \\
&[d_u + d_v]
\end{aligned}$$

Using Table 1, we get

$$\begin{aligned}
PM_1(Z) &= (N_1 + M + N_2)^{MN_1} \\
&\times (M + N_2 + N_1 + N_3)^{N_1N_2} \\
&\times (N_1 + N_3 + N_2 + N_4)^{N_2N_3} \times \dots \\
&+ (N_{r-3} + N_{r-1} + N_{r-2} + N_r)^{N_{r-2}N_{r-1}} \\
&\times (N_{r-2} + N_r + N_{r-1} + N)^{N_{r-1}N_r} \\
&\times (N_{r-1} + N + N_r)^{N_rN}. \tag{19}
\end{aligned}$$

This can be further summarized as follows:

$$\begin{aligned}
PM_1(Z) &= (N_1 + M + N_2)^{MN_1} \\
&\times (M + N_2 + N_1 + N_3)^{N_1N_2} \\
&\times \prod_{i=2}^{r-2} (N_{i-1} + N_{i+1} + N_i + N_{i+2})^{N_iN_{i+1}} \tag{20} \\
&\times (N_{r-2} + N_r + N_{r-1} + N)^{N_{r-1}N_r} \\
&\times (N_{r-1} + N + N_r)^{N_rN}.
\end{aligned}$$

$$\text{(iv) We know, from equation (6), } PM_2(Z) = \prod_{uv \in E} [d_u \times d_v]$$

$$\begin{aligned}
&= \prod_{uv \in E(N_1, M+N_2)} [d_u \times d_v] + \prod_{uv \in E(M+N_2, N_1+N_3)} [d_u \times d_v] \\
&+ \prod_{uv \in E(N_1+N_3, N_2+N_4)} [d_u \times d_v] + \dots \\
&+ \prod_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} [d_u \times d_v] \\
&+ \prod_{uv \in E(N_{r-2}+N_r, N_{r-1}+N)} [d_u \times d_v] \\
&+ \prod_{uv \in E(N_{r-1}+N, N_r)} [d_u \times d_v]. \tag{21}
\end{aligned}$$

Substituting the values from Table 1, we get

$$\begin{aligned}
PM_2(Z) &= [(N_1)(M+N_2)]^{MN_1} \\
&\times [(M+N_2)(N_1+N_3)]^{N_1N_2} \\
&\times [(N_1+N_3)(N_2+N_4)]^{N_2N_3} \times \dots \\
&\times [(N_{r-3}+N_{r-1})(N_{r-2}+N_r)]^{N_{r-2}N_{r-1}} \\
&\times [(N_{r-2}+N_r)(N_{r-1}+N)]^{N_{r-1}N_r} \\
&\times [(N_{r-1}+N)(N_r)]^{N_rN}. \tag{22}
\end{aligned}$$

The above expression can be expressed as follows:

$$\begin{aligned}
PM_2(Z) &= [(N_1)(M+N_2)]^{MN_1} \\
&\times [(M+N_2)(N_1+N_3)]^{N_1N_2} \\
&\times \prod_{i=2}^{r-2} [(N_{i-1}+N_{i+1})(N_i+N_{i+2})]^{N_iN_{i+1}} \tag{23} \\
&\times [(N_{r-2}+N_r)(N_{r-1}+N)]^{N_{r-1}N_r} \\
&\times [(N_{r-1}+N)(N_r)]^{N_rN}.
\end{aligned}$$

Theorem 3. Let $Z \cong DNN(N_1N_2 \dots N_r)$ be a deep neural network. Then the forgotten Zagreb index ($F(Z)$) and hyper-Zagreb index ($HM(Z)$) of DNN are given as follows:

$$\begin{aligned}
\text{(i) } F(Z) &= (MN_1)((N_1)^2 + (M+N_2)^2) + (N_1N_2) \\
&((M+N_2)^2 + (N_1+N_3)^2) + \sum_{i=2}^{r-2} N_iN_{i+1}((N_{i-1}+ \\
&N_{i+1})^2 + (N_i+N_{i+2})^2) + (N_{r-1}N_r)((N_{r-2}+N_r)^2 + \\
&(N_{r-1}+N)^2) + (N_rN)((N_{r-1}+N)^2 + (N_r)^2). \\
\text{(ii) } HM(Z) &= (MN_1)(N_1+M+N_2)^2 + (N_1N_2)(M+ \\
&N_2+N_1+N_3)^2 + \sum_{i=2}^{r-2} N_iN_{i+1}(N_{i-1}+N_{i+1}+N_i+ \\
&N_{i+2})^2 + (N_{r-1}N_r)(N_{r-2}+N_r+N_{r-1}+N)^2 + \\
&(N_rN)(N_{r-1}+N+N_r)^2.
\end{aligned}$$

Proof. To compute the topological indices of DNN, we use the edge partition method. In Table 1, we have calculated the degrees of end nodes of every edge for $DNN(N_1N_2 \dots N_r)$.

Now, by using the definitions and values from Table 1, we get the results given below

$$\begin{aligned}
(i) \quad F(Z) &= \sum_{uv \in E} [d_u^2 + d_v^2] = \sum_{uv \in E(N_1, M+N_2)} [d_u^2 + d_v^2] \\
&+ \sum_{uv \in E(M+N_2, N_1+N_3)} [d_u^2 + d_v^2] + \sum_{uv \in E(N_1+N_3, N_2+N_4)} \\
&[d_u^2 + d_v^2] + \dots + \sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} [d_u^2 + d_v^2] \\
&+ \sum_{uv \in E((N_{r-2}+N_r, N_{r-1}+N))} [d_u^2 + d_v^2] \\
&+ \sum_{uv \in E(N_{r-1}+N, N_r)} [d_u^2 + d_v^2].
\end{aligned}$$

Using Table 1, the above relation becomes

$$\begin{aligned}
F(Z) &= (MN_1)((N_1)^2 + (M+N_2)^2) \\
&+ (N_1N_2)((M+N_2)^2 + (N_1+N_3)^2) \\
&+ (N_2N_3)((N_1+N_3)^2 + (N_2+N_4)^2) \\
&+ \dots + (N_{r-2}N_{r-1})((N_{r-3}+N_{r-1})^2 + (N_{r-2}+N_r)^2) \\
&+ (N_{r-1}N_r)((N_{r-2}+N_r)^2 + (N_{r-1}+N)^2) \\
&+ (N_rN)((N_{r-1}+N)^2 + (N_r)^2).
\end{aligned} \tag{24}$$

This can be summarized as follows:

$$\begin{aligned}
F(Z) &= (MN_1)((N_1)^2 + (M+N_2)^2) \\
&+ (N_1N_2)((M+N_2)^2 + (N_1+N_3)^2) \\
&+ \sum_{i=2}^{r-2} N_i N_{i+1} ((N_{i-1}+N_{i+1})^2 + (N_i+N_{i+2})^2) \\
&+ (N_{r-1}N_r)((N_{r-2}+N_r)^2 + (N_{r-1}+N)^2) \\
&+ (N_rN)((N_{r-1}+N)^2 + (N_r)^2)
\end{aligned} \tag{25}$$

$$\begin{aligned}
(ii) \quad HM(Z) &= \sum_{uv \in E} (d_u + d_v)^2 = \sum_{uv \in E(N_1, M+N_2)} [d_u + d_v]^2 \\
&+ \sum_{uv \in E(M+N_2, N_1+N_3)} [d_u + d_v]^2 + \sum_{uv \in E(N_1+N_3, N_2+N_4)} \\
&[d_u + d_v]^2 + \dots + \sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} [d_u + d_v]^2 \\
&+ \sum_{uv \in E((N_{r-2}+N_r, N_{r-1}+N))} [d_u + d_v]^2 + \\
&\sum_{uv \in E(N_{r-1}+N, N_r)} [d_u + d_v]^2
\end{aligned}$$

Substituting values from Table 1, we get

$$\begin{aligned}
HM(Z) &= (MN_1)(N_1 + M + N_2)^2 \\
&+ (N_1N_2)(M + N_2 + N_1 + N_3)^2 \\
&+ (N_2N_3)(N_1 + N_3 + N_2 + N_4)^2 + \dots \\
&+ (N_{r-2}N_{r-1})(N_{r-3} + N_{r-1} + N_{r-2} + N_r)^2 \\
&+ (N_{r-1}N_r)(N_{r-2} + N_r + N_{r-1} + N)^2 \\
&+ (N_rN)(N_{r-1} + N + N_r)^2.
\end{aligned} \tag{26}$$

The above expression can be further summarized as follows:

$$\begin{aligned}
HM(Z) &= (MN_1)(N_1 + M + N_2)^2 \\
&+ (N_1N_2)(M + N_2 + N_1 + N_3)^2 \\
&+ \sum_{i=2}^{r-2} N_i N_{i+1} (N_{i-1} + N_{i+1} + N_i + N_{i+2})^2 \\
&+ (N_{r-1}N_r)(N_{r-2} + N_r + N_{r-1} + N)^2 \\
&+ (N_rN)(N_{r-1} + N + N_r)^2.
\end{aligned} \tag{27}$$

Theorem 4. Let $Z \cong DNN(N_1N_2 \dots N_r)$ be a deep neural network. The atom-bond connectivity index ($ABC(Z)$), geometric-arithmetic index ($GA(Z)$), sum connectivity index ($SCI(Z)$), and augmented Zagreb index ($AZI(Z)$) of DNN are given as follows:

$$\begin{aligned}
(i) \quad ABC(Z) &= (MN_1)\sqrt{(N_1 + M + N_2 - 2)/(N_1)(M+N_2)} \\
&+ (N_1N_2)\sqrt{(M + N_2 + N_1 + N_3 - 2)(M + N_2 + \\
&N_1 + N_3 - 2)/(M+N_2)(N_1 + N_3) + \sum_{i=2}^{r-2} N_i N_{i+1} \\
&+ (N_{r-1}N_r)\sqrt{(N_{r-2} + N_r + N_{r-1} + N - 2)/(N_{r-2} + N_r) \\
&(N_{r-1} + N) + \\
&(N_rN)\sqrt{((N_{r-1} + N + N_r - 2)/(N_{r-1} + N)(N_r))} \\
(ii) \quad GA(Z) &= 2(MN_1)\sqrt{((N_1)(M+N_2)/(N_1 + M+N_2))} + \\
&2(N_1N_2)\sqrt{((M+N_2)(N_1 + N_3)/(M+N_2 + N_1 + N_3))} + \\
&2 \sum_{i=2}^{r-2} N_i N_{i+1} \sqrt{((N_{i-1} + N_{i+1})(N_i + N_{i+2})/(N_{i-1} + N_{i+1} + \\
&N_i + N_{i+2}))} + 2(N_{r-1}N_r)\sqrt{((N_{r-2} + N_r)(N_{r-1} + \\
&N - 2)/(N_{r-2} + N_r + N_{r-1} + N))} + \\
&2(N_rN)\sqrt{((N_{r-1} + N)(N_r - 2)/(N_{r-1} + N + N_r))} \\
(iii) \quad SCI(Z) &= ((MN_1)/\sqrt{(N_1 + (M + N_2))}) + ((N_1N_2)/ \\
&\sqrt{(M + N_2) + (N_1 + N_3)}) + (\sum_{i=2}^{r-2} N_i N_{i+1} / \\
&((N_{i-1} + N_{i+1}) + (N_i + N_{i+2}))^{(1/2)}) + \\
&((N_{r-1}N_r)/\sqrt{(N_{r-2} + N_r)(N_{r-1} + N)}) + \\
&((N_rN)/\sqrt{(N_{r-1} + N) + (N_r)}) \\
(iv) \quad AZI(Z) &= (MN_1)((N_1)(M+N_2)/N_1 + M+N_2)^3 + \\
&(N_1N_2)((M+N_2)(N_1 + N_3)/M+N_2 + N_1 + N_3)^3 + \\
&\sum_{i=2}^{r-2} N_i N_{i+1} ((N_{i-1} + N_{i+1})(N_i + N_{i+2})/N_{i-1} + N_{i+1} +
\end{aligned}$$

$$\begin{aligned} & N_i + N_{i+2})^3 + (N_{r-1}N_r)((N_{r-2} + N_r)(N_{r-1} + N)/ \\ & N_{r-2} + N_r + N_{r-1} + N)^3 + (N_rN)((N_{r-1} + N) \\ & (N_r - 2)/N_{r-1} + N + N_r)^3 \end{aligned}$$

Proof

$$(i) \ ABC(Z) = \sum_{uv \in E} \sqrt{(d_u + d_v - 2/d_u d_v)} = \sum_{uv \in E(N_1, M+N_2)} \sqrt{(d_u + d_v - 2/d_u d_v)} +$$

$$\begin{aligned} & \sum_{uv \in E(M+N_2, N_1+N_3)} \sqrt{(d_u + d_v - 2/d_u d_v)} + \\ & \sum_{uv \in E(N_1+N_3, N_2+N_4)} \sqrt{(d_u + d_v - 2/d_u d_v)} + \dots + \\ & \sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} \sqrt{(d_u + d_v - 2/d_u d_v)} + \\ & \sum_{uv \in E(N_{r-2}+N_r, N_{r-1}+N)} \sqrt{(d_u + d_v - 2/d_u d_v)} + \\ & \sum_{uv \in E(N_{r-1}+N, N_r)} \sqrt{(d_u + d_v - 2/d_u d_v)} \end{aligned}$$

Using edge partition in Table 1, we have

$$\begin{aligned} ABC(Z) &= (MN_1) \sqrt{\frac{(N_1 + M + N_2 - 2)}{(N_1)(M+N_2)}} \\ &+ (N_1N_2) \sqrt{\frac{(M + N_2 + N_1 + N_3 - 2)}{(M+N_2)(N_1 + N_3)}} \\ &+ (N_2N_3) \sqrt{\frac{(N_1 + N_3 + N_2 + N_4 - 2)}{(N_1 + N_3)(N_2 + N_4)}} + \dots \\ &+ (N_{r-2}N_{r-1}) \sqrt{\frac{(N_{r-3} + N_{r-1} + N_{r-2} + N_r - 2)}{(N_{r-3} + N_{r-1})(N_{r-2} + N_r)}} \\ &+ (N_{r-1}N_r) \sqrt{\frac{(N_{r-2} + N_r + N_{r-1} + N - 2)}{(N_{r-2} + N_r)(N_{r-1} + N)}} \\ &+ (N_rN) \sqrt{\frac{(N_{r-1} + N + N_r - 2)}{(N_{r-1} + N)(N_r)}}. \end{aligned} \tag{28}$$

which can be shortened as follows:

$$\begin{aligned} ABC(Z) &= (MN_1) \sqrt{\frac{(N_1 + M + N_2 - 2)}{(N_1)(M+N_2)}} + (N_1N_2) \sqrt{\frac{(M + N_2 + N_1 + N_3 - 2)}{(M+N_2)(N_1 + N_3)}} \\ &+ \sum_{i=2}^{r-2} N_i N_{i+1} \sqrt{\frac{((N_{i-1} + N_{i+1} + N_i + N_{i+2} - 2)/(N_{i-1} + N_{i+1})(N_i + N_{i+2}))}{(N_{i-1} + N_{i+1})(N_i + N_{i+2})}} \\ &+ (N_{r-1}N_r) \sqrt{\frac{(N_{r-2} + N_r + N_{r-1} + N - 2)}{(N_{r-2} + N_r)(N_{r-1} + N)}} + (N_rN) \sqrt{\frac{(N_{r-1} + N + N_r - 2)}{(N_{r-1} + N)(N_r)}}. \end{aligned} \tag{29}$$

$$(ii) \quad GA(Z) = \sum_{uv \in E} (2\sqrt{d_u d_v} / d_u + d_v) = \sum_{uv \in E(N_1, M+N_2)} (2\sqrt{d_u d_v} / d_u + d_v) + \sum_{uv \in E(M+N_2, N_1+N_3)} (2\sqrt{d_u d_v} / d_u + d_v) + \sum_{uv \in E(N_1+N_3, N_2+N_4)} (2\sqrt{d_u d_v} / d_u + d_v) + \dots + \sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} (2\sqrt{d_u d_v} / d_u + d_v) +$$

$$\sum_{uv \in E((N_{r-2}+N_r, N_{r-1}+N))} (2\sqrt{d_u d_v} / d_u + d_v) + \sum_{uv \in E(N_{r-1}+N, N_r)} (2\sqrt{d_u d_v} / d_u + d_v)$$

Using Table 1, we get

$$GA(Z) = 2(MN_1) \sqrt{\frac{(N_1)(M+N_2)}{(N_1 + M+N_2)}} + 2(N_1N_2) \sqrt{\frac{(M+N_2)(N_1 + N_3)}{(M+N_2 + N_1 + N_3)}} + 2(N_2N_3) \sqrt{\frac{(N_1 + N_3)(N_2 + N_4)}{(N_1 + N_3 + N_2 + N_4)}} + \dots + 2(N_{r-2}N_{r-1}) \sqrt{\frac{(N_{r-3} + N_{r-1})(N_{r-2} + N_r)}{(N_{r-3} + N_{r-1} + N_{r-2} + N_r)}} + 2(N_{r-1}N_r) \sqrt{\frac{(N_{r-2} + N_r)(N_{r-1} + N)}{(N_{r-2} + N_r + N_{r-1} + N)}} + 2(N_rN) \sqrt{\frac{(N_{r-1} + N)(N_r)}{(N_{r-1} + N+N_r)}}.$$

This can be expressed as follows:

$$GA(Z) = 2(MN_1) \sqrt{\frac{(N_1)(M+N_2)}{(N_1 + M+N_2)}} + 2(N_1N_2) \sqrt{\frac{(M+N_2)(N_1 + N_3)}{(M+N_2 + N_1 + N_3)}} + 2 \sum_{i=2}^{r-2} N_i N_{i+1} \sqrt{\frac{((N_{i-1} + N_{i+1})(N_i + N_{i+2}))}{(N_{i-1} + N_{i+1} + N_i + N_{i+2})}} + 2(N_{r-1}N_r) \sqrt{\frac{(N_{r-2} + N_r)(N_{r-1} + N)}{(N_{r-2} + N_r + N_{r-1} + N)}} + 2(N_rN) \sqrt{\frac{(N_{r-1} + N)(N_r)}{(N_{r-1} + N+N_r)}}.$$

$$(iii) \quad SCI(Z) = \sum_{uv \in E} (1/\sqrt{d_u + d_v}) = \sum_{uv \in E(N_1, M+N_2)} (1/\sqrt{d_u + d_v}) + \sum_{uv \in E(M+N_2, N_1+N_3)} (1/\sqrt{d_u + d_v}) + \sum_{uv \in E(N_1+N_3, N_2+N_4)} (1/\sqrt{d_u + d_v}) + \dots + \sum_{uv \in E(N_{r-3}+N_{r-1}, N_{r-2}+N_r)} (1/\sqrt{d_u + d_v}) + \sum_{uv \in E(N_{r-2}+N_r, N_{r-1}+N)} (1/\sqrt{d_u + d_v}) + \sum_{uv \in E(N_{r-1}+N, N_r)} (1/\sqrt{d_u + d_v})$$

Using Table 1, we get

$$\begin{aligned}
 \text{SCI}(Z) &= \frac{(MN_1)}{\sqrt{(N_1) + (M + N_2)}} \\
 &+ \frac{(N_1N_2)}{\sqrt{(M + N_2) + (N_1+N_3)}} \\
 &+ \frac{(N_2N_3)}{\sqrt{(N_1 + N_3) + (N_2 + N_4)}} + \dots \\
 &+ \frac{(N_{r-2}N_{r-1})}{\sqrt{(N_{r-3}+N_{r-1}) + (N_{r-2}+N_r)}} \\
 &+ \frac{(N_{r-1}N_r)}{\sqrt{(N_{r-2} + N_r)(N_{r-1} + N)}} \\
 &+ \frac{(N_rN)}{\sqrt{(N_{r-1} + N) + (N_r)}}.
 \end{aligned}
 \tag{32}$$

This can be expressed as follows:

$$\begin{aligned}
 \text{SCI}(Z) &= \frac{(MN_1)}{\sqrt{(N_1) + (M + N_2)}} \\
 &+ \frac{(N_1N_2)}{\sqrt{(M + N_2) + (N_1+N_3)}} \\
 &+ \left(\sum_{i=2}^{r-2} N_i N_{i+1} / ((N_{i-1} + N_{i+1}) + (N_i + N_{i+2}))^{(1/2)} \right) \\
 &+ \frac{(N_{r-1}N_r)}{\sqrt{(N_{r-2} + N_r)(N_{r-1} + N)}} \\
 &+ \frac{(N_rN)}{\sqrt{(N_{r-1} + N) + (N_r)}}.
 \end{aligned}$$

$$\begin{aligned}
 \text{(iv) } AZI(Z) &= \sum_{uv \in E} (d_u \times d_v / d_u + d_v)^3 = \\
 &\sum_{uv \in E(N_1, M+N_2)} (d_u d_v / d_u + d_v)^3 + \\
 &\sum_{uv \in E(M+N_2, N_1+N_3)} (d_u d_v / d_u + d_v)^3 + \\
 &\sum_{uv \in E(N_1+N_3, N_2+N_4)} (d_u d_v / d_u + d_v)^3 + \dots + \\
 &\sum_{uv \in (N_{r-3}+N_{r-1}, N_{r-2}+N_r)} (d_u d_v / d_u + d_v)^3 + \\
 &\sum_{uv \in E((N_{r-2}+N_r, N_{r-1}+N))} (d_u d_v / d_u + d_v)^3 + \\
 &\sum_{uv \in E(N_{r-1}+N, N_r)} (d_u d_v / d_u + d_v)^3
 \end{aligned}$$

Substituting values from Table 1, we get

$$\begin{aligned}
 AZI(Z) &= (MN_1) \left(\frac{(N_1)(M+N_2)}{N_1 + M+N_2} \right)^3 \\
 &+ (N_1N_2) \left(\frac{(M+N_2)(N_1 + N_3)}{M+N_2 + N_1 + N_3} \right)^3 \\
 &+ (N_2N_3) \left(\frac{(N_1 + N_3)(N_2 + N_4)}{N_1 + N_3 + N_2 + N_4} \right)^3 + \dots \\
 &+ (N_{r-2}N_{r-1}) \left(\frac{(N_{r-3}+N_{r-1})(N_{r-2}+N_r - 2)}{N_{r-3} + N_{r-1} + N_{r-2} + N_r} \right)^3 \\
 &+ (N_{r-1}N_r) \left(\frac{(N_{r-2} + N_r)(N_{r-1} + N - 2)}{N_{r-2} + N_r + N_{r-1} + N} \right)^3 \\
 &+ (N_rN) \left(\frac{(N_{r-1} + N)(N_r - 2)}{N_{r-1} + N+N_r} \right)^3.
 \end{aligned}
 \tag{34}$$

This can be abbreviated as follows:

$$\begin{aligned}
 AZI(Z) &= (MN_1) \left(\frac{(N_1)(M+N_2)}{N_1 + M+N_2} \right)^3 \\
 &+ (N_1N_2) \left(\frac{(M+N_2)(N_1 + N_3)}{M+N_2 + N_1 + N_3} \right)^3 \\
 &+ \sum_{i=2}^{r-2} N_i N_{i+1} \left(\frac{(N_{i-1} + N_{i+1})(N_i + N_{i+2})}{N_{i-1} + N_{i+1} + N_i + N_{i+2}} \right)^3 \\
 &+ (N_{r-1}N_r) \left(\frac{(N_{r-2} + N_r)(N_{r-1} + N - 2)}{N_{r-2} + N_r + N_{r-1} + N} \right)^3 \\
 &+ (N_rN) \left(\frac{(N_{r-1} + N)(N_r - 2)}{N_{r-1} + N+N_r} \right)^3.
 \end{aligned}
 \tag{35}$$

4. Conclusions

The deep neural network is helpful in modeling compounds with desirable physical and chemical properties employing the structure of compounds. This paper gives computational insight into the degree-dependent topological indices, which include the Randic index, Zagreb index, multiplicative Zagreb indices, harmonic index, ABC index, GA index, and sum-connectivity index of a general DNN with r-hidden layers. These indices correlate the structure with the properties such as boiling point, molar refractivity (MR), molar volume (MV), polar surface area, surface tension, enthalpy of vaporization, flash point, and many others. The results

(33)

computed in the above theorems give generally closed formulas that can be exploited to compute the topological indices of neural networks under study by giving specific values to the input parameters. The values of the computed indices grow with the growth of hidden layers and also depend on the number of nodes in each layer.

A deep neural network is an important tool used in experimental design, data reduction, fault diagnosis, and process control. The QSAR studies must be integrated with the neural network approach in order to achieve a more physical understanding of the system. The use of DNN provides an alternative way of predicting physical properties and its linkage with topological indices can further enhance theoretical achievements.

This study can be extended further by analyzing the distance-based topological indices such as the Wiener index, Harary index, and PI index. Computation of spectral invariants of deep neural networks such as energy, Estrada energy, and Kirchhoff index is also open for further research in this area.

Data Availability

No data were used to support the findings of this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

The study was supported by the Science & Technology Bureau of Chengdu 2020-YF09-00005-SN and Sichuan Science and by the Technology program 2021YFH0107 Erasmus + SHYFTE Project 598649-EPP-1-2018-1-FR-EPPKA2-CBHE-JP and by the National Key Research and Development Program under Grant 2018YFB0904205.

References

- [1] J. Cao and R. Li, "Fixed-time synchronization of delayed memristor-based recurrent neural networks," *Science China Information Sciences*, vol. 60, no. 3, pp. 032201–032215, 2017.
- [2] T. P. Tran, T. T. S. Nguyen, P. Tsai, and X. Kong, "BSPNN: boosted subspace probabilistic neural network for email security," *Artificial Intelligence Review*, vol. 35, no. 4, pp. 369–382, 2011.
- [3] T. P. Tran, L. Cao, D. Tran, and C. D. Nguyen, "Novel intrusion detection using probabilistic neural network and adaptive boosting," *International Journal of Computer Science and Information Security*, vol. 6, no. 1, pp. 83–91, 2009.
- [4] M. Yang and H. Wei, "Application of a neural network for the prediction of crystallization kinetics," *Industrial & Engineering Chemistry Research*, vol. 45, no. 1, pp. 70–75, 2006.
- [5] O. S. Kharitonova, V. V. Bronskaya, T. V. Ignashina, A. A. Al-Muntaser, and L. E. Khairullina, "Modeling of absorption process using neural networks," *IOP Conference Series: Earth and Environmental Science*, vol. 315, no. 3, pp. 032025–032026, 2019.
- [6] A. Velásco-Mejía, V. Vallejo-Becerra, A. U. Chávez-Ramírez, J. Torres-González, Y. Reyes-Vidal, and F. Castañeda-Zaldívar, "Modeling and optimization of pharmaceutical crystallization process by using neural networks and genetic algorithms," *Powder Technology*, vol. 292, pp. 122–128, 2016.
- [7] M. Azzam, N. A. K. Aramouni, M. N. Ahmad, M. Awad, W. Kwapinski, and J. Zeaiter, "Dynamic optimization of dry reformer under catalyst sintering using neural networks," *Energy Conversion and Management*, vol. 157, pp. 146–156, 2018.
- [8] M. Bagheri, A. Akbari, and S. A. Mirbagheri, "Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: a critical review," *Process Safety and Environmental Protection*, vol. 123, pp. 229–252, 2019.
- [9] M. Randić, "Characterization of molecular branching," *Journal of the American Chemical Society*, vol. 97, no. 23, pp. 6609–6615, 1975.
- [10] B. Bollobas and P. Erdos, "Graphs of extremal weights," *Ars Combinatoria*, vol. 50, pp. 225–233, 1998.
- [11] I. Gutman and N. Trinajstić, "Graph theory and molecular orbitals. Total ϕ -electron energy of alternant hydrocarbons," *Chemical Physics Letters*, vol. 17, no. 4, pp. 535–538, 1972.
- [12] M. Ghorbani and M. Hosseinzadeh, "A new version of Zagreb indices," *Filomat*, vol. 26, no. 1, pp. 93–100, 2012.
- [13] G. H. Shirdel, H. Rezapour, and A. M. Sayadi, "The hyperzagreb index of graph operations," *Iranian Journal of Mathematical Chemistry*, vol. 4, pp. 213–220, 2013.
- [14] E. Estrada, L. Torres, L. Rodriguez, and I. Gutman, "An atom-bond connectivity index: modelling the enthalpy of formation of alkanes," *Indian Journal of Chemistry - Section A Inorganic, Physical, Theoretical and Analytical Chemistry*, vol. 37, pp. 849–855, 1998.
- [15] B. Zhou and N. Trinajstić, "On general sum-connectivity index," *Journal of Mathematical Chemistry*, vol. 47, no. 1, pp. 210–218, 2010.
- [16] D. Vukičević and B. Furtula, "Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges," *Journal of Mathematical Chemistry*, vol. 46, no. 4, pp. 1369–1376, 2009.
- [17] M. Javaid and J. Cao, "Computing topological indices of probabilistic neural network," *Neural Computing & Applications*, vol. 30, no. 12, pp. 3869–3876, 2018.
- [18] S. Mondal, N. De, and A. Pal, "Molecular descriptors of neural networks with chemical significance," *Revue Roumaine de Chimie*, vol. 65, no. 11, pp. 1031–1044, 2021.
- [19] J. B. Liu, Z. Raza, and M. Javaid, "Zagreb connection numbers for cellular neural networks," *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 8038304, 8 pages, 2020.
- [20] M. Javaid, M. Abbas, J. B. Liu, W. C. Teh, and J. Cao, "Topological properties of four-layered neural networks," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 2, pp. 111–122, 2019.
- [21] J. B. Liu, J. Zhao, S. Wang, M. Javaid, and J. Cao, "On the topological properties of the certain neural networks," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 8, no. 4, pp. 257–268, 2018.
- [22] C. Wang, J. B. Liu, and S. Wang, "Sharp upper bounds for multiplicative Zagreb indices of bipartite graphs with given diameter," *Discrete Applied Mathematics*, vol. 227, pp. 156–165, 2017.
- [23] R. Khalid, N. Idrees, and M. Jawwad Saif, "Topological characterization of Book graph and stacked Book graph,"

- Computers, Materials & Continua*, vol. 60, no. 1, pp. 41–54, 2019.
- [24] W. Gao, W. Wang, and M. R. Farahani, “Topological indices study of molecular structure in anticancer drugs,” *Journal of Chemistry*, vol. 2016, Article ID 3216327, 8 pages, 2016.
- [25] N. Idrees, M. Jawwad Saif, A. Sadiq, A. Rauf, and F. Hussain, “Topological indices of H-naphtalenic nanosheet,” *Open Chemistry*, vol. 16, no. 1, pp. 1184–1188, 2018.
- [26] N. Idrees, M. J. Saif, A. Rauf, and S. Mustafa, “First and second Zagreb eccentricity indices of thorny graphs,” *Symmetry*, vol. 9, no. 1, p. 7, 2017.
- [27] J. B. Liu, Y. Bao, W. T. Zheng, and S. Hayat, “Network coherence analysis on a family of nested weighted n-polygon networks,” *Fractals*, vol. 29, no. 08, pp. 2150260–2150276, 2021.
- [28] J. B. Liu, T. Zhang, Y. Wang, and W. Lin, “The Kirchhoff index and spanning trees of Möbius/cylinder octagonal chain,” *Discrete Applied Mathematics*, vol. 307, pp. 22–31, 2022.
- [29] S. Afridi, M. Yasin Khan, and G. Ali, “On generalized topological indices for some special graphs,” *Journal of Mathematics*, vol. 2022, p. 21, 2022.

Research Article

Computing the Entropy Measures for the Line Graphs of Some Chemical Networks

Muhammad Farhan Hanif ¹, Hasan Mahmood,^{1,2} Shazia Manzoor ³
and Fikre Bogale Petros ⁴

¹Abdus Salam School of Mathematical Sciences, Government College University, Lahore, Pakistan

²Department of Mathematics, Government College University, Lahore, Pakistan

³Department of Mathematics, COMSATS University Islamabad, Lahore Campus, Pakistan

⁴Department of Mathematics, Addis Ababa University, Addis Ababa, Ethiopia

Correspondence should be addressed to Fikre Bogale Petros; fikre.bogale@aau.edu.et

Received 29 June 2022; Accepted 18 August 2022; Published 6 October 2022

Academic Editor: Gohar Ali

Copyright © 2022 Muhammad Farhan Hanif 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.

Chemical Graph entropy plays a significant role to measure the complexity of chemical structures. It has explicit chemical uses in chemistry, biology, and information sciences. A molecular structure of a compound consists of many atoms. Especially, the hydrocarbons is a chemical compound that consists of carbon and hydrogen atoms. In this article, we discussed the concept of subdivision of chemical graphs and their corresponding line chemical graphs. More precisely, we discuss the properties of chemical graph entropies and then constructed the chemical structures namely triangular benzenoid, hexagonal parallelogram, and zigzag edge coronoid fused with starphene. Also, we estimated the degree-based entropies with the help of line graphs of the subdivision of above mentioned chemical graphs.

1. Introduction

Mathematical chemistry is a field of theoretical chemistry that uses mathematical approaches to discuss molecular structure without necessarily referring to quantum mechanics [1]. Chemical Graph Theory is a branch of mathematical chemistry where a chemical phenomenon is theoretically described using graph theory [2, 3]. The growth of organic disciplines has been aided by Chemical Graph Theory [4, 5]. In mathematical chemistry, graph invariants or topological indices are numeric quantities that describe various essential features of organic components and are produced from an analogous molecular graph [6, 7]. Degree-based indices are among the topological indices used to predict bioactivity, boiling point, draining energy, stability, and physico-chemical properties of certain chemical compounds [8, 9]. Due to their chemical applications, these indices have significant role in theoretical chemistry. Zhang et al. [10–12] discuss the

topological indices of generalized bridge molecular graphs, Carbon Nanotubes and product of chemical graphs. Zhang et al. [13–15] provided the physical analysis of heat for formation and entropy of Ceria Oxide. For further study about indices, see [16, 17]. Shannon [18] originated the conception of information entropy in communication theory. However, it was later discovered as a quantity that applied to all things with a set nature [19, 20], including molecular graphs [21–23]. In chemistry, information entropy is now used in two modes. Firstly, it is a structural descriptor for assessing the complexity of chemical structures [24]. Information entropy is useful in this regard for connecting structural and physico-chemical features [25], numerically distinguishing isomers of organic molecules [26], and classifying natural products and synthetic chemicals [27, 28]. The physico-chemical sounding of information entropy is a different mode of application. As a result, Terenteva and Kobozev demonstrated its utility in analyzing physico-chemical processes that simulate

information transmission [29]. Zhdanov [30] used entropy values to study organic compound chemical processes. The information entropy is defined as:

$$\begin{aligned} ENT_{\psi}(\mathcal{F}) &= -\sum_{i=1}^q N_i \frac{\Lambda(l_i m_i)}{In} \log \frac{\Lambda(l_i m_i)}{In}, \\ &= \log(In) - \frac{1}{In} \sum_{i=1}^q N_i \Lambda(l_i m_i) \log \Lambda(l_i m_i). \end{aligned} \quad (1)$$

Here, the logarithm is considered to be with base e while \mathcal{F}_V , \mathcal{F}_E and $\Lambda(lm)$ represent the vertex set, the edge set and the edge weight of the edge (lm) in Λ . Many graph entropies have been calculated in the literature utilising characteristic polynomials, vertices degree, and graph order [31–34]. Graph entropies, which are based on independent sets, matchings, and the degree of vertices [35], have been estimated in recent years. Dehmer and Mowshowits proposed several graph complexity and Hosoya entropy relationships [23, 32, 36, 37]. For further study, see [19, 21, 38–42, 59, 60]. The graph \mathcal{F} is structured into ordered pairs, with one object being referred to as a vertex set (\mathcal{F}_V) and the other as an edge set (\mathcal{F}_E), and these vertices and edges being connected. When two vertices of \mathcal{F} share an edge, they are said to be neighboring. The sum of the degrees of all neighboring vertices of l is denoted by A_l , and the degree of a vertex l is represented by $\hat{\mathcal{N}}(l)$. By replacing each of $S(\mathcal{F})$'s edges with a path of length two, the subdivision graph $S(\mathcal{F})$ is formed. The line graph is denoted by the symbol $L(\mathcal{F})$ in which $|V(L(\mathcal{F}))| = |E(\mathcal{F})|$ and two vertices of $L(\mathcal{F})$ are adjacent iff their corresponding edges share a common end points in \mathcal{F} .

1.1. Randić Entropy [43, 44]. If $\Lambda(lm) = (\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m))^\alpha$, with $\alpha = 1, -1, 1/2, -1/2$, then

$$\sum_{lm \in \mathcal{F}_E} \Lambda(lm) = \sum_{lm \in \mathcal{F}_E} (\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m))^\alpha = R_\alpha. \quad (2)$$

Now (1) represent the Randic Entropy.

$$\begin{aligned} ENT_{R_\alpha}(\mathcal{F}) &= \log(R_\alpha) - \frac{1}{(R_\alpha)} \sum_{i=1}^q \sum_{lm \in \mathcal{F}_E} [(\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m))^\alpha] \\ &\quad \cdot \log [(\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m))^\alpha]. \end{aligned} \quad (3)$$

1.2. Atom Bond Connectivity Entropy [45]. If $\Lambda(lm) = \sqrt{\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m) - 2/\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m)}$, then

$$\sum_{lm \in \mathcal{F}_E} \Lambda(lm) = \sum_{lm \in \mathcal{F}_E} \sqrt{\frac{\hat{\mathcal{N}}(l) + \hat{\mathcal{N}}(m) - 2}{\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m)}} = ABC(\mathcal{F}). \quad (4)$$

Thus (1) is converted in the following form:

$$\begin{aligned} ENT_{ABC}(\mathcal{F}) &= \log(ABC) \\ &\quad - \frac{1}{(ABC)} \sum_{i=1}^q \sum_{lm \in \mathcal{F}_E} \left[\sqrt{\frac{\hat{\mathcal{N}}(l) + \hat{\mathcal{N}}(m) - 2}{\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m)}} \right] \\ &\quad \cdot \log \left[\sqrt{\frac{\hat{\mathcal{N}}(l) + \hat{\mathcal{N}}(m) - 2}{\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m)}} \right]. \end{aligned} \quad (5)$$

1.3. The Geometric Arithmetic Entropy [43, 44]. If $\Lambda(lm) = 2\sqrt{\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m)}/\hat{\mathcal{N}}(l) + \hat{\mathcal{N}}(m)$, then

$$\sum_{lm \in \mathcal{F}_E} \Lambda(lm) = \sum_{lm \in \mathcal{F}_E} \frac{2\sqrt{\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m)}}{\hat{\mathcal{N}}(l) + \hat{\mathcal{N}}(m)} = GA(\mathcal{F}). \quad (6)$$

Now (1) takes the form as given below.

$$\begin{aligned} ENT_{GA}(\mathcal{F}) &= \log(GA) \\ &\quad - \frac{1}{(GA)} \sum_{i=1}^q \sum_{lm \in \mathcal{F}_E} \left[\frac{2\sqrt{\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m)}}{\hat{\mathcal{N}}(l) + \hat{\mathcal{N}}(m)} \right] \\ &\quad \cdot \log \left[\frac{2\sqrt{\hat{\mathcal{N}}(l) \times \hat{\mathcal{N}}(m)}}{\hat{\mathcal{N}}(l) + \hat{\mathcal{N}}(m)} \right]. \end{aligned} \quad (7)$$

1.4. The Fourth Atom Bond Connectivity Entropy [35]. If $\Lambda(lm) = \sqrt{A_l + A_m - 2/A_l A_m}$, then

$$\sum_{lm \in E(\mathcal{F})} \Lambda(lm) = \sum_{lm \in E(\mathcal{F})} \sqrt{\frac{A_l + A_m - 2}{A_l A_m}} = ABC_4(\mathcal{F}). \quad (8)$$

Now (1) converted in the following form as:

$$\begin{aligned} ENT_{ABC_4}(\mathcal{F}) &= \log(ABC_4(\mathcal{F})) - \frac{1}{(ABC_4(\mathcal{F}))} \sum_{i=1}^q \sum_{lm \in E_i(\mathcal{F})} \\ &\quad \cdot \log \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right]^{\left[\sqrt{A_l + A_m - 2/A_l A_m} \right]}. \end{aligned} \quad (9)$$

1.5. The Fifth Geometric Arithmetic Entropy [35]. If $\Lambda(lm) = 2\sqrt{A_l A_m}/A_l + A_m$, then

$$\sum_{lm \in E(\mathcal{F})} \Lambda(lm) = \sum_{lm \in E(\mathcal{F})} \frac{2\sqrt{A_l A_m}}{A_l + A_m} = GA_5(\mathcal{F}). \quad (10)$$

Equation (1) is now changed to the following form, which is known as fifth geometric arithmetic entropy.

$$ENT_{GA_5}(\mathcal{F}) = \log(GA_5(\mathcal{F})) - \frac{1}{(GA_5(\mathcal{F}))} \sum_{i=1}^q \sum_{lm \in E_i(\mathcal{F})} \left[\log \left[\frac{2\sqrt{A_l A_m}}{A_l + A_m} \right] \right]. \quad (11)$$

See [35, 44] for further information on these entropy measures.

2. Formation of Triangular Benzenoid $T_x \forall x \in \mathbb{N}$

Triangular benzenoids are a group of benzenoid molecular graphs and are denoted by T_x , where x characterizes the number of hexagons at the bottom of the graph and $1/2x(x+1)$ represents the total number of hexagons in T_x . Triangular benzenoids are a generalization of the benzene molecule C_6H_6 , with benzene rings forming a triangular shape. In physics, chemistry, and nanosciences, the benzene molecule is a common molecule. Synthesizing aromatic chemicals is quite fruitful [46]. Raut [47] calculated some topological indices for the triangular benzenoid system. Hussain et al. [48] discussed the irregularity determinants of some benzenoid systems.

Kwun [49] calculated degree-based indices by using M polynomials. For further details, see [50, 51]. The hexagons are placed in rows, with each row increasing by one hexagon. For T_1 , there are only one type of edges $e_1 = (2, 2)$ and $|e_1| = 6$. Therefore, $V(T_1) = 6$ and $E(T_1) = 6$ while three kinds of edges are there in T_2 e.g. $e_1 = (2, 2)$, $e_2 = (2, 3)$, $e_3 = (3, 3)$ and $|e_1| = 6$, $|e_2| = 6$, $|e_3| = 3$. Therefore, $V(T_2) = 13$ and $E(T_2) = 15$. Continuing in this way, $|V(T_x)| = x^2 + 4x + 1$ and $|E(T_x)| = 3/2x(x+3)$. The subdivision graph of T_x and its line graph are demonstrated in Figure 1. It is to be noted that $|V(L(S(T_x)))| = 3x(x+3)$ and $|E(L(S(T_x)))| = 3/2(3x^2 + 7x - 2)$.

Let $\mathcal{F} = L(S(T_x))$. i.e. \mathcal{F} is the line graph of the subdivision graph of triangular benzenoid T_x . We will use the edge partition and vertices counting technique to compute our abstracted indices and entropies. The degree of each edge's terminal vertices is used in the edge partitioning of \mathcal{F} . It is easy to see that there are only three types of edges shown in Table 1.

2.1. Entropy Measure for $L(S(T_x))$. We'll calculate the entropies of $\mathcal{F} = L(S(T_x))$ in this section.

2.1.1. Randic Entropy of $L(S(T_x))$. The Randic index and entropy for $\alpha = 1, -1, 1/2, -1/2$, with the help of Table 1, and equation (3) is:

$$ENT_{R_\alpha}(\mathcal{F}) = \log(R_\alpha) - \frac{1}{(R_\alpha)} \sum_{i=1}^3 \sum_{lm \in E_i(\mathcal{F})} [(\hat{N}(l) \times \hat{N}(m))^\alpha] \log [(\hat{N}(l) \times \hat{N}(m))^\alpha] \quad (12)$$

$$= \log(R_\alpha) - \frac{1}{(R_\alpha)} \left[[4^\alpha(3x+9) \times \log(4^\alpha)] + [6^\alpha(6x-6) \times \log(6^\alpha)] + \left[\frac{3^{(2\alpha+1)}}{2} (3x^2+x-4) \times \log(9^\alpha) \right] \right].$$

By putting $\alpha = 1, -1, 1/2, -1/2$, in (3), we get the Randic entropies as given below:

$$ENT_{R_1}(\mathcal{F}) = \log\left(\frac{3}{2}(27x^2 + 41x - 52)\right) - \frac{12(x+3) \times \log[4]}{(3/2)(27x^2 + 41x - 52)} - \frac{36(x-1) \times \log[6]}{(3/2)(27x^2 + 41x - 52)} - \frac{27/2(3x^2+x-4) \times \log[9]}{(3/2)(27x^2 + 41x - 52)}$$

$$ENT_{R_{-1}}(\mathcal{F}) = \log\left(\frac{1}{12}(6x^2 + 23x + 7)\right) + \frac{3/4(x+3) \times \log[4]}{(1/12)(6x^2 + 23x + 7)} + \frac{(x-1) \times \log[6]}{(1/12)(6x^2 + 23x + 7)} + \frac{1/6(3x^2+x-4) \times \log[9]}{(1/12)(6x^2 + 23x + 7)},$$

$$ENT_{R_{1/2}}(\mathcal{F}) = \log\left(\frac{3}{2}(9x^2 + (7+4\sqrt{6})x - 4\sqrt{6})\right) - \frac{6(x+3) \times \log[2]}{(3/2)(9x^2 + (7+4\sqrt{6})x - 4\sqrt{6})} - \frac{6\sqrt{6}(x-1) \times \log[\sqrt{6}]}{(3/2)(9x^2 + (7+4\sqrt{6})x - 4\sqrt{6})}$$

$$- \frac{9/2(3x^2+x-4) \times \log[3]}{(3/2)(9x^2 + (7+4\sqrt{6})x - 4\sqrt{6})},$$

$$ENT_{R_{-1/2}}(\mathcal{F}) = \log\left(\frac{1}{2}(3x^2 + 2(2+\sqrt{6})x + 5)\right) + \frac{3/2(x+3) \times \log[2]}{(1/2)(3x^2 + 2(2+\sqrt{6})x + 5)} + \frac{\sqrt{6}(x-1) \times \log[\sqrt{6}]}{(1/2)(3x^2 + 2(2+\sqrt{6})x + 5)}$$

$$+ \frac{1/2(3x^2+x-4) \times \log[3]}{(1/2)(3x^2 + 2(2+\sqrt{6})x + 5)}.$$

(13)

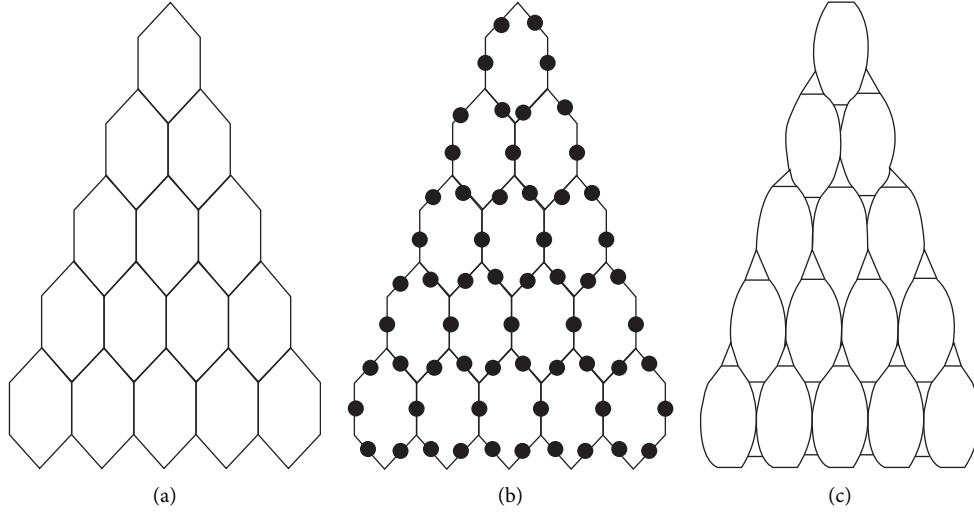


FIGURE 1: (a) Triangular benzenoid T_5 , (b) Subdivision of T_5 , (c) The line graph of subdivision graph of T_5 .

TABLE 1: Edge partition of $L(S(T_x))$.

$(\hat{N}(l), \hat{N}(m))$	N_i	Set of Edges
(2, 2)	$2(x + 3)$	E_1
(2, 3)	$6(x - 1)$	E_2
(3, 3)	$3/2(3x^2 + x - 4)$	E_3

2.1.2. *The ABC Entropy of $L(S(T_x))$.* The ABC index and entropy measure with the help of Table 1 and equation (5) is:

$$\begin{aligned}
 \text{ABC}(\mathcal{F}) &= 3x^2 + \left(\frac{9}{\sqrt{2}} + 1\right)x + \frac{3}{\sqrt{2}} - 4, \\
 \text{ENT}_{\text{ABC}}(\mathcal{F}) &= \log(\text{ABC}) - \frac{1}{(\text{ABC})} \sum_{i=1}^3 \sum_{lm \in E_i(\mathcal{F})} \left[\sqrt{\frac{\hat{N}(l) + \hat{N}(m) - 2}{\hat{N}(l) \times \hat{N}(m)}} \right] \log \left[\sqrt{\frac{\hat{N}(l) + \hat{N}(m) - 2}{\hat{N}(l) \times \hat{N}(m)}} \right] \\
 &= \log \left(3x^2 + \left(\frac{9}{\sqrt{2}} + 1\right)x + \frac{3}{\sqrt{2}} - 4 \right) + \frac{1/\sqrt{2}(9x + 3) \times \log[\sqrt{2}]}{(3x^2 + (9/\sqrt{2} + 1)x + 3/\sqrt{2} - 4)} - \frac{(3x^2 + x - 4) \times \log[2/3]}{(3x^2 + (9/\sqrt{2} + 1)x + 3/\sqrt{2} - 4)}.
 \end{aligned} \tag{14}$$

2.1.3. *The Geometric Arithmetic Entropy of $L(S(T_x))$.*

The GA index and entropy measure with the help of Table 1 and equation (7) is:

$$\begin{aligned}
 \text{GA}(\mathcal{F}) &= \frac{9}{2}x^2 + \frac{3x}{10}(8\sqrt{6} + 15) - \frac{3}{5}(4\sqrt{6} - 5), \\
 \text{ENT}_{\text{GA}}(\mathcal{F}) &= \log(\text{GA}) - \frac{1}{(\text{GA})} \sum_{i=1}^3 \sum_{lm \in E_i(\mathcal{F})} \left[\frac{2\sqrt{\hat{N}(l) \times \hat{N}(m)}}{\hat{N}(l) + \hat{N}(m)} \right] \log \left[\frac{2\sqrt{\hat{N}(l) \times \hat{N}(m)}}{\hat{N}(l) + \hat{N}(m)} \right] \\
 &= \log \left(\frac{9}{2}x^2 + \left(\frac{24\sqrt{6} + 45}{10}\right)x + \frac{15 - 12\sqrt{6}}{5} \right) - \frac{12\sqrt{6}/5(x - 1) \times \log[2\sqrt{6}/5]}{(9/2x^2 + (24\sqrt{6} + 45/10)x + 15 - 12\sqrt{6}/5)}.
 \end{aligned} \tag{15}$$

TABLE 2: Edge partition of $L(S(T_x))$.

(A_l, A_m)	N_i	Set of Edges
(4, 4)	9	\mathcal{F}_{E_1}
(4, 5)	6	\mathcal{F}_{E_2}
(5, 5)	$3(x-2)$	\mathcal{F}_{E_3}
(5, 8)	$6(y-1)$	\mathcal{F}_{E_4}
(8, 8)	$3(x-1)$	\mathcal{F}_{E_5}
(8, 9)	$6(x-1)$	\mathcal{F}_{E_6}
(9, 9)	$3/2(3x^2 + 2 - 5x)$	\mathcal{F}_{E_7}

2.1.4. *The ABC_4 Entropy of $L(S(T_x))$.* The edge partition of the graph $L(S(T_x))$ is grounded on the degree addition of terminal vertices of every edge, as shown in Table 2.

$$\begin{aligned}
 ENT_{ABC_4}(\mathcal{F}) &= \log(ABC_4) - \frac{1}{(ABC_4)} \sum_{i=1}^7 \sum_{lm \in E_i(\mathcal{F})} \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right] \log \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right], \\
 ENT_{ABC_4}(\mathcal{F}) &= \log(ABC_4) - \frac{[3\sqrt{6}/2] \log [\sqrt{6}/4]}{(ABC_4)} - \frac{[3\sqrt{7}/\sqrt{5}] \log [\sqrt{7}/2\sqrt{5}]}{(ABC_4)} \\
 &\quad - \frac{[6\sqrt{2}/5](x-1) \log [2\sqrt{2}/5]}{(ABC_4)} - \frac{[3\sqrt{11}/\sqrt{10}](x-1) \log [\sqrt{11}/2\sqrt{10}]}{(ABC_4)} \\
 &\quad - \frac{[3\sqrt{14}/8](x-1) \log [\sqrt{14}/8]}{(ABC_4)} - \frac{[\sqrt{15}/\sqrt{2}](x-1) \log [\sqrt{15}/6\sqrt{2}]}{(ABC_4)} - \frac{2/3(3x^2 - 5x + 2) \log [4/9]}{(ABC_4)}.
 \end{aligned} \tag{17}$$

If we consider $x = 1$, Then $ABC_4(\mathcal{F}) = 9\sqrt{6}/4$, and $ENT_{ABC_4}(\mathcal{F}) = 2.1972$.

2.1.5. *The GA_5 Entropy of $L(S(T_x))$.* After some simple calculations, the GA_5 index may be calculated using Table 2 under the constraint that $x \neq 1$.

$$\begin{aligned}
 ENT_{GA_5}(\mathcal{F}) &= \log(GA_5) - \frac{1}{(GA_5)} \sum_{i=1}^7 \sum_{lm \in E_i(\mathcal{F})} \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right] \log \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right], \\
 ENT_{GA_5}(\mathcal{F}) &= \log(GA_5) - \frac{[8\sqrt{5}/3] \log [4\sqrt{5}/9]}{(GA_5)} - \frac{[24\sqrt{10}/13](x-1) \log [4\sqrt{10}/13]}{(GA_5)} - \frac{[72\sqrt{2}/17](x-1) \log [12\sqrt{2}/17]}{(GA_5)}.
 \end{aligned} \tag{19}$$

3. Formation of Hexagonal Parallelogram

Nanotubes $H(x, y)$, $\forall x, y \in \mathbb{N}$

Hexagonal parallelogram nanotubes are formed by arranging hexagons in a parallelogram fashion. Baig et al. [52] computed counting polynomials of benzoid carbon nanotubes. Also, see [53]. We will denote this structure by $H(x, y) \forall x, y \in \mathbb{N}$, in which x and y represent the quantity of hexagons in any row and column respectively. Also, the order and size of $H(x, y)$ is $2(x + y + xy)$ and $3xy + 2x +$

After simple calculations, by using Table 2 subject to the condition that $x \neq 1$, we get

$$\begin{aligned}
 ABC_4(\mathcal{F}) &= \frac{3\sqrt{6}}{2} + \frac{3\sqrt{7}}{\sqrt{5}} + \frac{6\sqrt{2}}{5}(x-2) \\
 &\quad + \left(\frac{3\sqrt{11}}{\sqrt{10}} + \frac{3\sqrt{14}}{8} + \frac{\sqrt{15}}{\sqrt{2}} \right)(x-1) \\
 &\quad + \frac{2}{3}(3x^2 - 5x + 2).
 \end{aligned} \tag{16}$$

By using (9), the ABC_4 entropy as follows:

$$\begin{aligned}
 GA_5(\mathcal{F}) &= 3 + \frac{8\sqrt{5}}{3} + 3x + \left(\frac{24\sqrt{10}}{13} + \frac{72\sqrt{2}}{17} + 3 \right)(x-1) \\
 &\quad + \frac{3}{2}(3x^2 - 5n + 2).
 \end{aligned} \tag{18}$$

Therefore, (11), with Table 2 converted in the form:

$2y - 1$ respectively. The subdivision graph of $H(x, y)$ and its line graph is shown in Figure 2, see [46]. Let $\mathcal{F} = L(S(H(x, y)))$, then $|\mathcal{F}_V| = 2(3xy + 2x + 2y - 1)$ and $|\mathcal{F}_E| = 9xy + 4x + 4y - 5$. To compute our results, we will use edge partition technique which is grounded on the degree of terminal vertices of every edge. It is to be noted that there are only three types of edges, see Figure 2. The edge partition of chemical graph $L(S(H(x, y)))$ depending on the degree of terminal vertices is presented in Table 3.

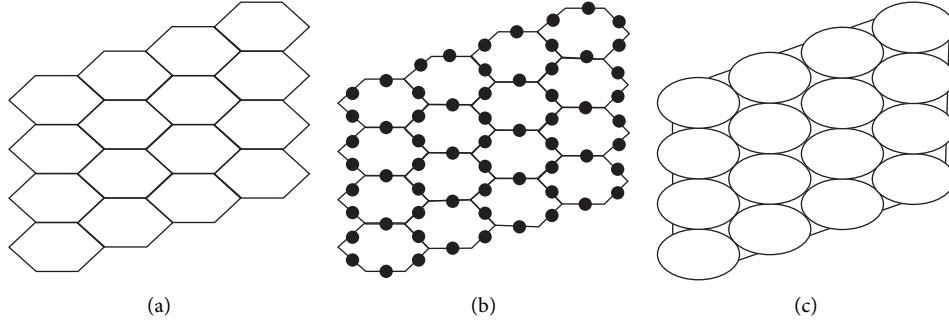


FIGURE 2: (a) Hexagonal parallelogram $H(x, y)$, (b) Subdivision of $H(x, y)$, (c) The line graph of subdivision graph of $H(x, y)$.

TABLE 3: Edge partition of $L(S(H(x, y)))$.

$(\widehat{\mathcal{N}}(l), \widehat{\mathcal{N}}(m))$	N_i	Kinds of Edges
(2, 2)	$2(4 + y + x)$	\mathcal{F}_{E_1}
(2, 3)	$4(-2 + y + x)$	\mathcal{F}_{E_2}
(3, 3)	$9xy - 2m - 2n - 5$	\mathcal{F}_{E_3}

3.1. Entropy Measure for $L(S(H(x, y)))$. We will enumerate the entropies of $\mathcal{F} = L(S(H(x, y)))$ in this section.

3.1.1. Randić Entropy of \mathcal{F} . The Randić index for $\alpha = 1, -1, 1/2, -1/2$, by using Table 3 is:

$$R_\alpha(\mathcal{F}) = 2(x + y + 4) \times (4)^\alpha + 4(x + y - 2) \times (6)^\alpha + (9xy - 2x - 2y - 5) \times (9)^\alpha. \quad (20)$$

So the (3) with Table 3 gives the Randić entropy and is converted in the form:

$$\begin{aligned} ENT_{R_\alpha}(\mathcal{F}) &= \log(R_\alpha) - \frac{1}{(R_\alpha)} \sum_{i=1}^3 \sum_{lm \in E_i(\mathcal{F})} [(\widehat{\mathcal{N}}(l) \times \widehat{\mathcal{N}}(m))^\alpha] \log [(\widehat{\mathcal{N}}(l) \times \widehat{\mathcal{N}}(m))^\alpha] \\ &= \log(R_\alpha) - \frac{1}{(R_\alpha)} \left[[4^\alpha (2x + 2y + 8) \times \log(4^\alpha)] + t[6^\alpha (4x + 4y - 8) \times \log(6^\alpha)] + n + q[9^\alpha (9xy - 2x - 2y - 5) \times \log(9^\alpha)] \right]. \end{aligned} \quad (21)$$

Now substitute $\alpha = 1, -1, 1/2, -1/2$, in (20), we get the Randić entropies as given below:

$$ENT_{R_1}(\mathcal{F}) = \log(81xy + 14(x + y) - 61) - \frac{8(x + y + 4) \times [4]}{(81xy + 14(x + y) - 61)} - \frac{24(x + y - 2) \times \log[6]}{(81xy + 14(x + y) - 61)} - \frac{9(9xy - 2x - 2y - 5) \times \log[9]}{(81xy + 14(x + y) - 61)}.$$

$$\begin{aligned} ENT_{R_{-1}}(\mathcal{F}) &= \log\left(xy + \frac{17}{18}(x + y) + \frac{1}{9}\right) + \frac{1/2(x + y + 4) \times [4]}{(xy + 17/18(x + y) + 1/9)} + \frac{2/3(x + y - 2) \times \log[6]}{(xy + 17/18(x + y) + 1/9)} \\ &\quad + \frac{1/9(9xy - 2x - 2y - 5) \times \log[9]}{(xy + 17/18(x + y) + 1/9)}. \end{aligned}$$

$$\begin{aligned} ENT_{R_{1/2}}(\mathcal{F}) &= \log(27xy + (4\sqrt{6} - 2)(x + y) + 1 - 8\sqrt{6}) - \frac{4(x + y + 4) \times [2]}{(27xy + (4\sqrt{6} - 2)(x + y) + 1 - 8\sqrt{6})} \\ &\quad - \frac{4\sqrt{6}(x + y - 2) \times \log[\sqrt{6}]}{(27xy + (4\sqrt{6} - 2)(x + y) + 1 - 8\sqrt{6})} - \frac{93(9xy - 2x - 2y - 5) \times \log[3]}{(27xy + (4\sqrt{6} - 2)(x + y) + 1 - 8\sqrt{6})}, \end{aligned}$$

$$\begin{aligned} ENT_{R_{-1/2}}(\mathcal{F}) &= \log\left(3xy + \left(\frac{1}{3} + \frac{4}{\sqrt{6}}\right)(x + y) + \frac{7}{3} - \frac{8}{\sqrt{6}}\right) + \frac{(x + y + 4) \times \log[2]}{(3xy + (1/3 + 4/\sqrt{6})(x + y) + 7/3 - 8/\sqrt{6})} \\ &\quad + \frac{4/\sqrt{6}(x + y - 2) \times \log[\sqrt{6}]}{(3xy + (1/3 + 4/\sqrt{6})(x + y) + 7/3 - 8/\sqrt{6})} + \frac{1/3(9xy - 2x - 2y - 5) \times \log[9]}{(3xy + (1/3 + 4/\sqrt{6})(x + y) + 7/3 - 8/\sqrt{6})}. \end{aligned} \quad (22)$$

3.1.2. *The ABC Entropy of \mathcal{F}* . With the use of Table 3 and equation (5), we can calculate the ABC index and entropy measure as follows:

$$ABC(\mathcal{F}) = 6xy + \left(\frac{9\sqrt{2} - 4}{3}\right)(x + y) - \frac{10}{3}. \quad (23)$$

Therefore, the equation (5), with Table 3 becomes as following and is called the atom bond connectivity entropy.

$$\begin{aligned} ENT_{ABC}(\mathcal{F}) &= \log(ABC) \\ &- \frac{1}{(ABC)} \sum_{i=1}^3 \sum_{lm \in E_i(\mathcal{F})} \left[\sqrt{\frac{\hat{N}(l) + \hat{N}(m) - 2}{\hat{N}(l) \times \hat{N}(m)}} \right] \\ &\cdot \log \left[\sqrt{\frac{\hat{N}(l) + \hat{N}(m) - 2}{\hat{N}(l) \times \hat{N}(m)}} \right], \\ &= \log \left(6xy + \left(\frac{9\sqrt{2} - 4}{3}\right)(x + y) - \frac{10}{3} \right) \\ &+ \frac{\sqrt{2}(x + y + 4) \times [\sqrt{2}]}{(6xy + (9\sqrt{2} - 4/3)(x + y) - 10/3)} \\ &+ \frac{2\sqrt{2}(x + y - 2) \times \log[\sqrt{6}]}{(6xy + (9\sqrt{2} - 4/3)(x + y) - 10/3)} \\ &- \frac{2/3(9xy - 2x - 2y - 5) \times \log[2/3]}{(6xy + (9\sqrt{2} - 4/3)(x + y) - 10/3)}. \end{aligned} \quad (24)$$

3.1.3. *The Geometric Arithmetic Entropy of \mathcal{F}* . We can calculate the GA index and entropy measure using Table 3 and equation (7) as follows:

$$GA(\mathcal{F}) = \frac{1}{5}(45xy + 8\sqrt{6}(x + y) + 15 - 16\sqrt{6}),$$

$$\begin{aligned} ENT_{GA}(\mathcal{F}) &= \log \left(\frac{1}{5}(45xy + 8\sqrt{6}(x + y) + 15 - 16\sqrt{6}) \right) \\ &- \frac{8\sqrt{6}/5(x + y - 2) \times \log[2\sqrt{6}/5]}{(1/5(45xy + 8\sqrt{6}(x + y) + 15 - 16\sqrt{6}))}. \end{aligned} \quad (25)$$

3.1.4. *The ABC_4 Entropy of \mathcal{F}*

Case 1. when $x > 1, y \neq 1$

The edge partition of $L(S(H(x, y)))$ is shown in Table 4.

Therefore, the ABC_4 index and entropy measure with the help of Table 4 and equation (9) yield as:

TABLE 4: Edge partition of $L(S(H(x, y)))$.

(A_l, A_m)	N_i	Kinds of edges
(4, 4)	8	\mathcal{F}_{E_1}
(4, 5)	8	\mathcal{F}_{E_2}
(5, 5)	$2(-4 + y + x)$	\mathcal{F}_{E_3}
(5, 8)	$4(-2 + y + x)$	\mathcal{F}_{E_4}
(8, 8)	$2(-2 + x + y)$	\mathcal{F}_{E_5}
(8, 9)	$2(-2 + x + y)$	\mathcal{F}_{E_6}
(9, 9)	$9xy - 8x - 8y + 7$	\mathcal{F}_{E_7}

$$\begin{aligned} ABC_4(\mathcal{F}) &= 4xy + \left(\frac{4\sqrt{2}}{5} + \frac{2\sqrt{11}}{\sqrt{10}} + \frac{\sqrt{14}}{4} + \frac{\sqrt{30}}{3} - \frac{32}{9}\right) \\ &\cdot (x + y) + 2\sqrt{6} + \frac{4\sqrt{7}}{\sqrt{5}} - \frac{16\sqrt{2}}{5} - \frac{4\sqrt{11}}{\sqrt{10}} \\ &- \frac{\sqrt{14}}{2} - \frac{2\sqrt{30}}{3} + \frac{28}{9}. \end{aligned} \quad (26)$$

Since \mathcal{F} has seven kinds of edges, So (9) by using Table 4 is converted in the form:

$$\begin{aligned} ENT_{ABC_4}(\mathcal{F}) &= \log(ABC_4) \\ &- \frac{1}{(ABC_4)} \sum_{i=1}^7 \sum_{lm \in E_i(\mathcal{F})} \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right] \\ &\cdot \log \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right], \\ ENT_{ABC_4}(\mathcal{F}) &= \log(ABC_4) - \frac{2\sqrt{6} \log[\sqrt{6}/4]}{(ABC_4)} \\ &- \frac{4\sqrt{7}/\sqrt{5} \log[\sqrt{7}/2\sqrt{5}]}{(ABC_4)} \\ &- \frac{4\sqrt{2}/5(x + y - 4) \log[2\sqrt{2}/5]}{(ABC_4)} \\ &- \frac{2\sqrt{11}/\sqrt{10}(x + y - 2) \log[\sqrt{11}/2\sqrt{10}]}{(ABC_4)} \\ &- \frac{\sqrt{14}/4(x + y - 2) \log[\sqrt{14}/8]}{(ABC_4)} \\ &- \frac{2\sqrt{15}/3\sqrt{2}(x + y - 2) \log[\sqrt{15}/6\sqrt{2}]}{(ABC_4)} \\ &- \frac{4/9(9xy - 8x - 8y + 7) \log[4/9]}{(ABC_4)}. \end{aligned} \quad (27)$$

Case 2. when $x = 1, y \neq 1$

By using the same process, we get the closed expressions for the ABC_4 index and ABC_4 entropy as:

$$\begin{aligned}
BC_4(\mathcal{F}) &= \left(\frac{4\sqrt{2}}{5} + \frac{2\sqrt{11}}{\sqrt{10}} + \frac{\sqrt{14}}{4} + \frac{\sqrt{30}}{3} + \frac{4}{9} \right) y + \frac{5\sqrt{6}}{2} \\
&\quad + \frac{2\sqrt{7}}{\sqrt{5}} - \frac{8\sqrt{2}}{5} - \frac{2\sqrt{11}}{\sqrt{10}} - \frac{\sqrt{30}}{3} - \frac{\sqrt{14}}{4} - A \frac{9}{4}, \\
ENT_{ABC_4}(\mathcal{F}) &= \log(ABC_4) - \frac{5\sqrt{6}/2 \log[\sqrt{6}/4]}{(ABC_4)} \\
&\quad - \frac{2\sqrt{7}/\sqrt{5} \log[\sqrt{7}/2\sqrt{5}]}{(ABC_4)} \\
&\quad - \frac{4\sqrt{2}/5(y-2) \log[2\sqrt{2}/5]}{(ABC_4)} \\
&\quad - \frac{2\sqrt{11}/\sqrt{10}(y-1) \log[\sqrt{11}/2\sqrt{10}]}{(ABC_4)} \\
&\quad - \frac{\sqrt{14}/4(y-1) \log[\sqrt{14}/8]}{(ABC_4)} \\
&\quad - \frac{2\sqrt{15}/3\sqrt{2}(y-1) \log[\sqrt{15}/6\sqrt{2}]}{(ABC_4)} \\
&\quad - \frac{4/9(y-1) \log[4/9]}{(ABC_4)}. \tag{28}
\end{aligned}$$

3.1.5. The Fifth Geometric Arithmetic Entropy of \mathcal{F}

Case 3. when $x > 1$, $y \neq 1$ The fifth geometric arithmetic entropy can be estimated by using (11), and Table 4 in the following manner:

$$\begin{aligned}
GA_5(\mathcal{F}) &= 9xy + \left(\frac{16\sqrt{10}}{13} + \frac{48\sqrt{2}}{17} - 4 \right) (x+y) + 3 \\
&\quad + \frac{32\sqrt{5}}{9} - \frac{32\sqrt{10}}{13} - \frac{96\sqrt{2}}{17}. \tag{29}
\end{aligned}$$

So the (11), with Table 4 can be written as:

$$\begin{aligned}
ENT_{GA_5}(\mathcal{F}) &= \log(GA_5) - \frac{1}{(GA_5)} \sum_{i=1}^7 \sum_{lm \in E_i(\mathcal{F})} \\
&\quad \cdot \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right] \log \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right] \\
&= \log(GA_5) - \frac{32\sqrt{5}/9 \log[4\sqrt{5}/9]}{(GA_5)} \tag{30} \\
&\quad - \frac{16\sqrt{10}/13(x+y-2) \log[4\sqrt{10}/13]}{(GA_5)} \\
&\quad - \frac{48\sqrt{2}/17(x+y-2) \log[12\sqrt{2}/17]}{(GA_5)}.
\end{aligned}$$

TABLE 5: Edge partition of $L(S(H(x, y)))$, for $x = 1$.

(A_l, A_m)	N_i	Kinds of edges
(4, 4)	10	\mathcal{F}_{E_1}
(4, 5)	4	\mathcal{F}_{E_2}
(5, 5)	$2(y-2)$	\mathcal{F}_{E_3}
(5, 8)	$4(y-1)$	\mathcal{F}_{E_4}
(8, 8)	$2(y-1)$	\mathcal{F}_{E_5}
(8, 9)	$2(y-1)$	\mathcal{F}_{E_6}
(9, 9)	$y-1$	\mathcal{F}_{E_7}

Case 4. when $x = 1$, $y \neq 1$ By using Table 5 and using (11) we get the closed expressions for the GA_5 index and GA_5 entropy as:

$$\begin{aligned}
GA_5(\mathcal{F}) &= \left(5 + \frac{16\sqrt{10}}{13} + \frac{48\sqrt{2}}{17} \right) y + 3 + \frac{16\sqrt{5}}{9} \\
&\quad - \frac{16\sqrt{10}}{13} - \frac{48\sqrt{2}}{17} - \frac{16\sqrt{10}}{13},
\end{aligned}$$

$$\begin{aligned}
ENT_{GA_5}(\mathcal{F}) &= \log(GA_5) - \frac{16\sqrt{5}/9 \log[4\sqrt{5}/9]}{(GA_5)} \tag{31} \\
&\quad - \frac{16\sqrt{10}/13(y-1) \log[4\sqrt{10}/13]}{(GA_5)} \\
&\quad - \frac{48\sqrt{2}/17(y-1) \log[12\sqrt{2}/17]}{(GA_5)}.
\end{aligned}$$

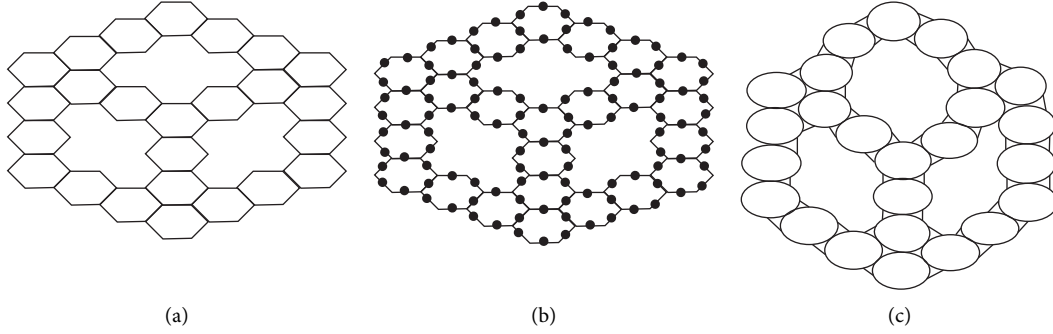
4. Formation from Fusion of Zigzag-Edge Coronoid with Starphene ZCS(x, y, z) Nanotubes

If a zigzag-edge coronoid $ZC(x, y, z)$ is fused with a starphene $St(x, y, z)$, then we will obtain a composite benzenoid. It is to be noted that $|V(ZCS(x, y, z))| = 36x - 54$ and $|E(ZCS(x, y, z))| = -63 + 15(z + y + x)$. The subdivision graph of $ZCS(x, y, z)$ and its line graph are illustrated in Figure 3. We can see from figures that the order and the size in the line graph of the subdivision graph of $ZCS(x, y, z)$ are $-126 + 30(z + y + x)$ and $-153 + 39(z + y + x)$ respectively [46]. Let \mathcal{F} represents the subdivision graph of $ZCS(x, y, z)$'s line graph. The edge division is determined by the degree of each edge's terminal vertices. Table 6 illustrates this.

4.1. Entropy Measure for $L(S(ZCS(x, y, z)))$. We'll calculate the entropies of $\mathcal{F} = L(S(ZCS(x, y, z)))$ in this section.

4.1.1. Randić Entropy of \mathcal{F} . For $\alpha = 1, -1, 1/2, -1/2$, the Randić index with the help of Table 1 is

$$\begin{aligned}
R_\alpha(\mathcal{F}) &= 6(x + y + z - 5) \times (4)^\alpha + 12(x + y + z - 7) \\
&\quad \times (6)^\alpha + (21x + 21y + 21z - 39) \times (9)^\alpha. \tag{32}
\end{aligned}$$

FIGURE 3: (a) $ZCS(4, 4, 4)$, (b) subdivision of $ZCS(4, 4, 4)$, (c) $L(S(ZCS(4, 4, 4)))$.TABLE 6: Edge partition of $L(S(ZCS))$.

$(\hat{N}(l), \hat{N}(m))$	N_i	Kinds of Edges
(2, 2)	$6(-5 + z + y + x)$	\mathcal{F}_{E_1}
(2, 3)	$12(-7 + z + y + x)$	\mathcal{F}_{E_2}
(3, 3)	$-39 + 21(z + y + x)$	\mathcal{F}_{E_3}

Using (3) Randić entropy is:

$$\begin{aligned}
 ENT_{R_\alpha}(\mathcal{F}) &= \log(R_\alpha) - \frac{1}{(R_\alpha)} \sum_{i=1}^3 \sum_{lm \in E_i(\mathcal{F})} \left[(\hat{N}(l) \times \hat{N}(m)^\alpha) \log [(\hat{N}(l) \times \hat{N}(m)^\alpha)] \right. \\
 &= \log(R_\alpha) - \frac{1}{(R_\alpha)} \left[[4^\alpha (6(x + y + z - 5)) \times \log(4^\alpha)] + [6^\alpha (12(x + y + z - 7)) \times \log(6^\alpha)] \right. \\
 &\quad \left. \left. + [(21(x + y + z) - 39) \times \log(9^\alpha)] \right] \right].
 \end{aligned} \tag{33}$$

By putting $\alpha = 1, -1, 1/2, -1/2$, in (32), we get the Randić entropies as given below:

$$\begin{aligned}
 ENT_{R_1}(\mathcal{F}) &= \log(-975 + 285(z + y + x)) - \frac{24(-5 + z + y + x) \times \log[4]}{(-975 + 285(z + y + x))} - \frac{72(x + y + z - 7) \times \log[6]}{(-975 + 285(z + y + x))} \\
 &\quad - \frac{189(z + tyn + qx) - t351) \times \log[9]}{(-975 + 285(z + y + x))}, \\
 ENT_{R_{-1}}(\mathcal{F}) &= \log\left(-\frac{131}{6} + \frac{35}{6}(z + y + x)\right) + \frac{3/2(-5 + z + y + x) \times \log[4]}{(-131/6 + 35/6(z + y + x))} + \frac{2(-7 + z + y + x) \times \log[6]}{(-131/6 + 35/6(z + y + x))} \\
 &\quad + \frac{1/3(-13 + 7(z + y + x)) \times \log[9]}{(-131/6 + 35/6(z + y + x))}, \\
 ENT_{R_{1/2}}(\mathcal{F}) &= \log((3 + 12\sqrt{6})(x + y + z) - 177 - 84\sqrt{6}) - \frac{12(x + y + z - 5) \times \log[2]}{((3 + 12\sqrt{6})(x + y + z) - 177 - 84\sqrt{6})} \\
 &\quad - \frac{6\sqrt{6}(x + y + z - 7) \times \log[\sqrt{6}]}{((3 + 12\sqrt{6})(x + y + z) - 177 - 84\sqrt{6})} - \frac{9(7(x + y + z) - 13) \times \log[3]}{((3 + 12\sqrt{6})(x + y + z) - 177 - 84\sqrt{6})}, \\
 ENT_{R_{-1/2}}(\mathcal{F}) &= \log((10 + 2\sqrt{6})(x + y + z) - 28 - 14\sqrt{6}) + \frac{3(x + y + z - 5) \times \log[2]}{((10 + 2\sqrt{6})(x + y + z) - 28 - 14\sqrt{6})} \\
 &\quad + \frac{2\sqrt{6}(x + y + z - 7) \times \log[\sqrt{6}]}{((10 + 2\sqrt{6})(x + y + z) - 28 - 14\sqrt{6})} + \frac{(7(x + y + z) - 13) \times \log[3]}{((10 + 2\sqrt{6})(x + y + z) - 28 - 14\sqrt{6})}.
 \end{aligned} \tag{34}$$

4.1.2. *The ABC Entropy of \mathcal{F}* . The ABC index and entropy measure with the help of Table 6 and equation (5) are:

$$\begin{aligned}
 ABC(\mathcal{F}) &= (14 + 9\sqrt{2})(x + y + z) - 26 - 57\sqrt{2}, \\
 ENT_{ABC}(\mathcal{F}) &= \log((14 + 9\sqrt{2})(x + y + z) - 26 - 57\sqrt{2}) \\
 &\quad + \frac{3\sqrt{2}(3(x + y + z) - 19) \times \log[\sqrt{2}]}{((14 + 9\sqrt{2})(x + y + z) - 26 - 57\sqrt{2})} \\
 &\quad - \frac{(14(x + y + z) - 26)\log[2/3]}{((14 + 9\sqrt{2})(x + y + z) - 26 - 57\sqrt{2})}. \tag{35}
 \end{aligned}$$

TABLE 7: Edge partition of $L(S(ZCS(x, y, z)))$ established on degree sum of terminal vertices, for every $x = y = z \geq 4$

(A_l, A_m)	N_i	Kinds of Edges
(4, 4)	6	\mathcal{F}_{E_1}
(4, 5)	12	\mathcal{F}_{E_2}
(5, 5)	$6(x + y + z - 8)$	\mathcal{F}_{E_3}
(5, 8)	$12(x + y + z - 7)$	\mathcal{F}_{E_4}
(8, 8)	$6(x + y + z - 9)$	\mathcal{F}_{E_5}
(8, 9)	$12(x + y + z - 5)$	\mathcal{F}_{E_6}
(9, 9)	$3(x + y + z + 25)$	\mathcal{F}_{E_7}

4.1.3. *The Geometric Arithmetic Entropy of \mathcal{F}* . The GA index and corresponding entropy with the help of Table 6 and equation (7) are:

$$GA(\mathcal{F}) = (x + y + z)(27 + 24\sqrt{6}/5) - 69 - 168\sqrt{6}/5,$$

$$ENT_{GA}(\mathcal{F}) = \log((x + y + z)(27 + 24\sqrt{6}/5) - 69 - 168\sqrt{6}/5) - \frac{24\sqrt{6}/5(x + y + z - 7) \times \log[2\sqrt{6}/5]}{(x + y + z)(27 + 24\sqrt{6}/5) - 69 - 168\sqrt{6}/5}. \tag{36}$$

4.1.4. *The ABC_4 entropy of \mathcal{F}* . Table 7 shows the graph $L(S(ZCS(x, y, z)))$'s edge partition, which is based on the degree addition of each edge's terminal vertices.

After simple calculations, the ABC_4 index and entropy measure with the help of Table 7 and equation (9) subject to the condition that $x = y = z \geq 4$

$$ABC_4(\mathcal{F}) = (x + y + z) \left(\sqrt{30} + \frac{4}{3} + \frac{12\sqrt{2}}{5} + \frac{3\sqrt{14}}{4} + \frac{6\sqrt{11}}{\sqrt{10}} \right) - 5\sqrt{30} + \frac{100}{3} - \frac{96\sqrt{2}}{5} + \frac{27\sqrt{14}}{4} - \frac{42\sqrt{11}}{\sqrt{10}} + \frac{3\sqrt{6}}{2} + \frac{6\sqrt{7}}{\sqrt{5}},$$

$$ENT_{ABC_4}(\mathcal{F}) = \log(ABC_4) - \frac{1}{(ABC_4)} \sum_{i=1}^7 \sum_{lm \in E_i(\mathcal{F})} \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right] \log \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right],$$

$$\begin{aligned}
 ENT_{ABC_4}(\mathcal{F}) &= \log(ABC_4) - \frac{[3\sqrt{6}/2]\log[\sqrt{6}/4]}{(ABC_4)} - \frac{[12\sqrt{2}/5](x + y + z - 8)\log[2\sqrt{2}/5]}{(ABC_4)} \\
 &\quad - \frac{[6\sqrt{11}/\sqrt{10}](x + y + z - 7)\log[\sqrt{11}/2\sqrt{10}]}{(ABC_4)} - \frac{3\sqrt{14}/4(x + y + z - 9)\log[\sqrt{14}/8]}{(ABC_4)} \\
 &\quad - \frac{\sqrt{30}(x + y + z - 5)\log[\sqrt{15}/6\sqrt{2}]}{(ABC_4)} - \frac{4/3(x + y + z + 25)\log[4/9]}{(ABC_4)} - \frac{[6\sqrt{7}/\sqrt{5}]\log[\sqrt{7}/2\sqrt{5}]}{(ABC_4)}. \tag{37}
 \end{aligned}$$

TABLE 8: Comparison of randic entropies for $L(S(T_x))$.

$[x]$	ENT_{R_1}	$ENT_{R_{-1}}$	$ENT_{R_{1/2}}$	$ENT_{R_{-1/2}}$
[46]	0.4055	2.5590	2.4849	2.6263
[52]	3.1863	3.0463	3.5667	3.5970
[25]	4.0316	3.6767	4.2203	4.2280
[26]	4.5797	4.2928	4.6981	4.6991
[24]	4.9945	4.8714	5.0779	5.0764
[23]	5.3312	5.4107	5.3942	5.3918
[27]	5.6159	5.9131	5.6658	5.6631
[2]	5.8632	6.3820	5.9041	5.9013
[56]	6.0820	6.8208	6.1164	6.1136
[31]	6.2785	7.2325	6.3080	6.3053

TABLE 9: Comparison of ENT_{ABC} , ENT_{GA} , ENT_{ABC_4} , and ENT_{GA_5} for $L(S(T_x))$.

$[x]$	ENT_{ABC}	ENT_{GA}	ENT_{ABC_4}	ENT_{GA_5}
[46]	2.3116	2.4849	2.1972	0
[52]	3.5239	3.5835	3.5749	3.5835
[25]	4.2025	4.2341	4.2263	4.2341
[26]	4.6897	4.7095	4.7028	4.7095
[24]	5.0739	5.0876	5.0817	5.0876
[23]	5.3926	5.4027	5.3975	5.4026
[27]	5.6655	5.6733	5.6687	5.6733
[2]	5.9046	5.91087	5.9066	5.9108
[56]	6.1174	6.1225	6.1187	6.1225
[31]	6.3093	6.3135	6.3100	6.3135

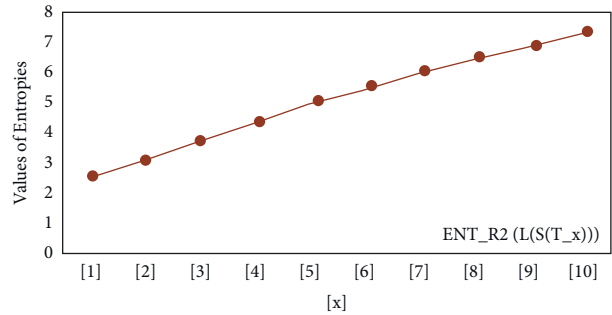
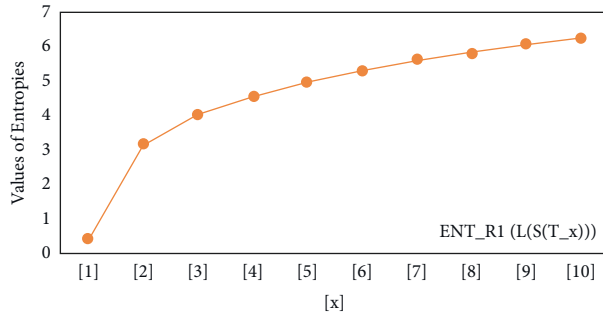


FIGURE 4: (a) R_1 entropy, (b) R_{-1} entropy.

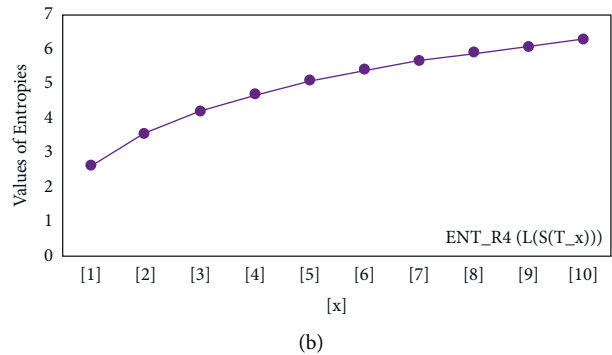
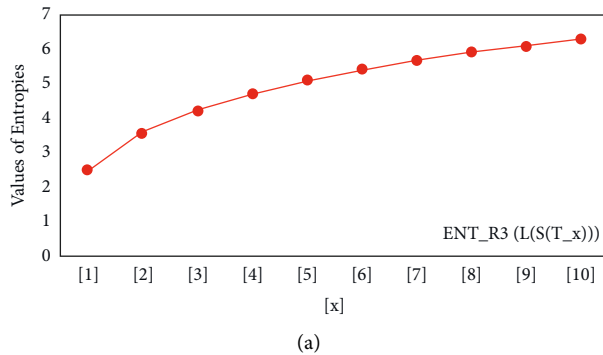


FIGURE 5: (a) $R_{1/2}$ entropy, (b) $R_{-1/2}$ entropy.

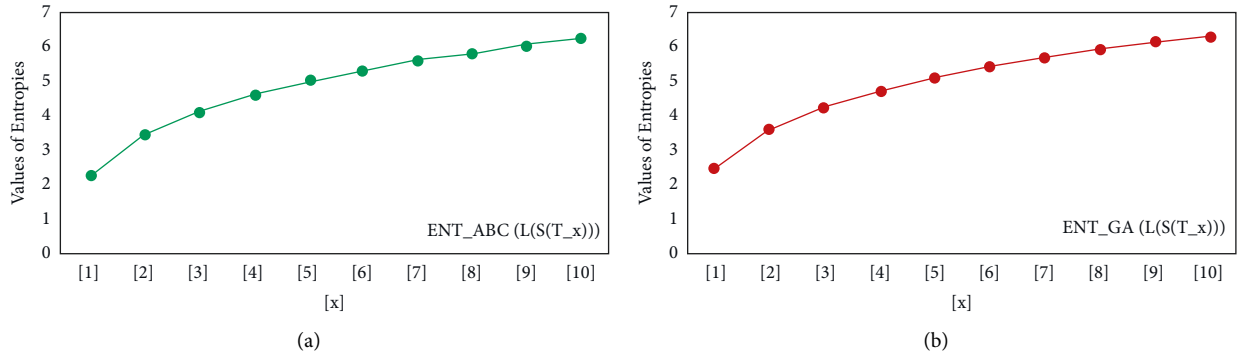


FIGURE 6: (a) The ABC entropy, (b) The GA entropy.

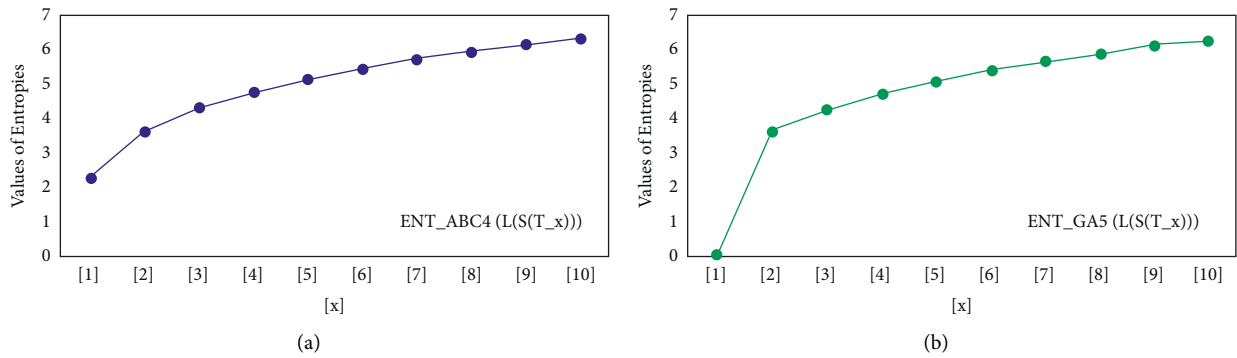


FIGURE 7: (a) The ABC_4 entropy, (b) The GA_5 entropy.

TABLE 10: Comparison of randic entropies for $L(S(H(x, y)))$.

$[x, y]$	ENT_{R_1}	$ENT_{R_{-1}}$	$ENT_{R_{1/2}}$	$ENT_{R_{-1/2}}$
[1, 1]	2.4849	2.4849	2.4849	2.4849
[2, 2]	3.7917	3.7830	3.8344	3.8332
[3, 3]	4.5635	4.5428	4.5933	4.5906
[4, 4]	5.1096	5.0872	5.1323	5.1294
[5, 5]	5.5345	5.5129	5.5530	5.5502
[6, 6]	5.8833	5.8630	5.8988	5.8962
[7, 7]	6.1794	6.1615	6.1928	6.1904
[8, 8]	6.4368	6.4194	6.4486	6.4464
[9, 9]	6.6646	6.6483	6.6751	6.6731
[10, 10]	6.8688	6.5370	6.8783	6.8822

TABLE 12: Comparison of ENT_{ABC_4} and ENT_{GA_5} Entropies for $L(S(H(x, y)))$, $x > 1$ and $y \neq 1$.

$[x, y]$	ENT_{ABC_4}	ENT_{GA_5}
[2, 2]	3.7879	3.4822
[3, 3]	4.5387	2.2596
[4, 4]	5.0783	4.8387
[5, 5]	5.5018	5.2952
[6, 6]	5.8509	5.6704
[7, 7]	6.1481	5.9882
[8, 8]	6.4068	6.2636
[9, 9]	6.6360	6.5064
[10, 10]	6.8417	6.7234

TABLE 11: Comparison of ENT_{ABC} and ENT_{GA} entropies for $L(S(H(x, y)))$.

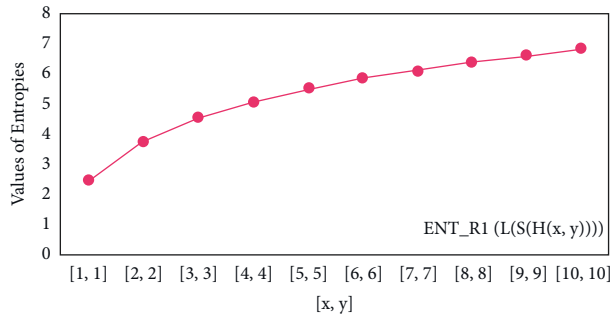
$[x, y]$	ENT_{ABC}	ENT_{GA}
[1, 1]	2.4849	2.4849
[2, 2]	3.8497	3.8501
[3, 3]	4.6048	4.6051
[4, 4]	5.1413	5.1416
[5, 5]	5.5604	5.5607
[6, 6]	5.9051	5.9053
[7, 7]	6.1982	6.1985
[8, 8]	6.4534	6.4536
[9, 9]	6.6794	6.6796
[10, 10]	6.8822	6.8824

TABLE 13: Comparison of ENT_{ABC_4} and ENT_{GA_5} entropies for $L(S(H(x, y)))$, $x = 1$ and $y \neq 1$.

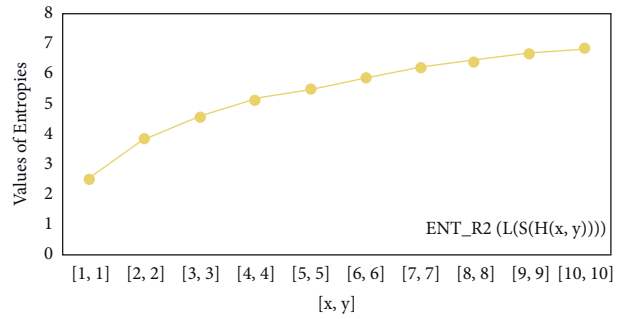
$[y]$	ENT_{ABC_4}	ENT_{GA_5}
[52]	3.1846	3.2958
[25]	3.5933	3.6888
[26]	3.8884	3.9702
[24]	4.1184	4.1896
[23]	4.3064	4.3694
[27]	4.4653	4.5217
[2]	4.6027	4.6539
[56]	4.7238	4.7706
[31]	4.8319	4.8751

TABLE 14: Comparison of randic entropies for $L(S(ZCS(x, y, z)))$.

$[x, y, z]$	ENT_{R_1}	$ENT_{R_{-1}}$	$ENT_{R_{1/2}}$	$ENT_{R_{-1/2}}$
[4, 4, 4]	5.7200	5.70060	5.7432	5.7407
[5, 5, 5]	6.0342	6.0165	6.0587	6.0565
[6, 6, 6]	6.2730	6.2564	6.2982	6.2961
[7, 7, 7]	6.4657	6.4497	6.4913	6.4893
[8, 8, 8]	6.6272	6.6117	6.6531	6.6511
[9, 9, 9]	6.7662	6.7511	6.7923	6.7904
[10, 10, 10]	6.8883	6.8734	6.9145	6.9126

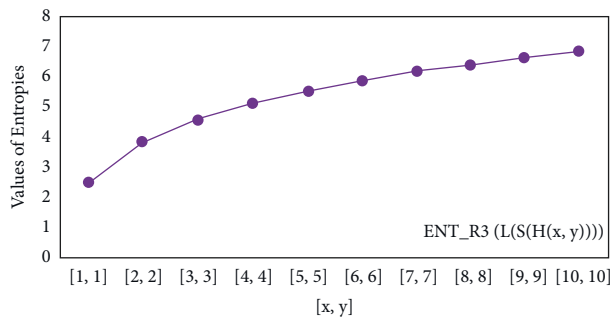


(a)

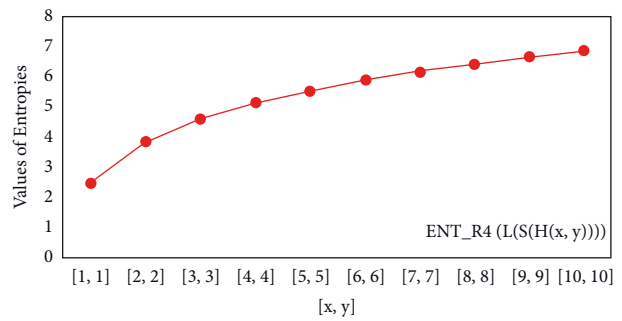


(b)

FIGURE 8: (a) R_1 entropy, (b) R_{-1} entropy.

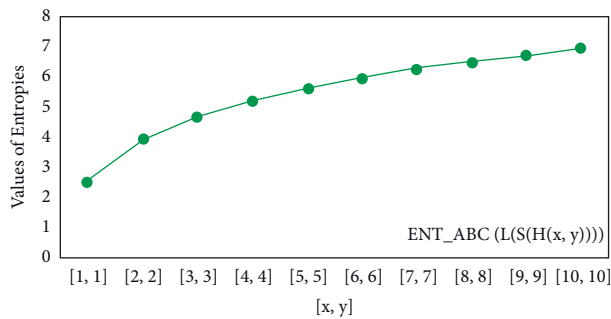


(a)

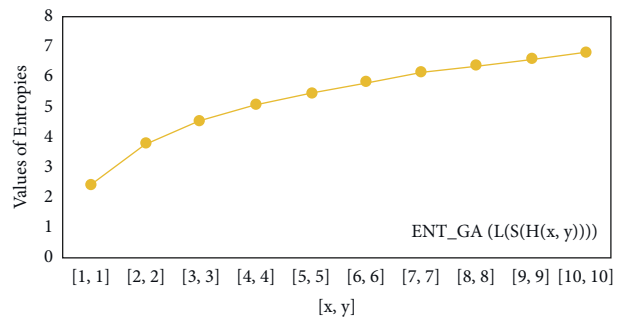


(b)

FIGURE 9: (a) $R_{1/2}$ entropy, (b) $R_{-1/2}$ entropy.



(a)



(b)

FIGURE 10: (a) The ABC entropy, (b) The GA, entropy.

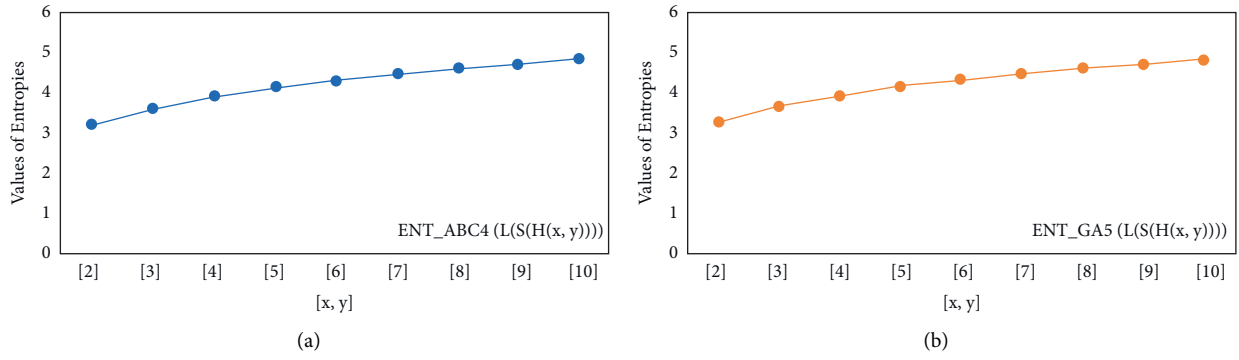


FIGURE 11: (a) The ABC_4 entropy, (b) The GA_5 entropy, $x \geq 1, y \neq 1$.

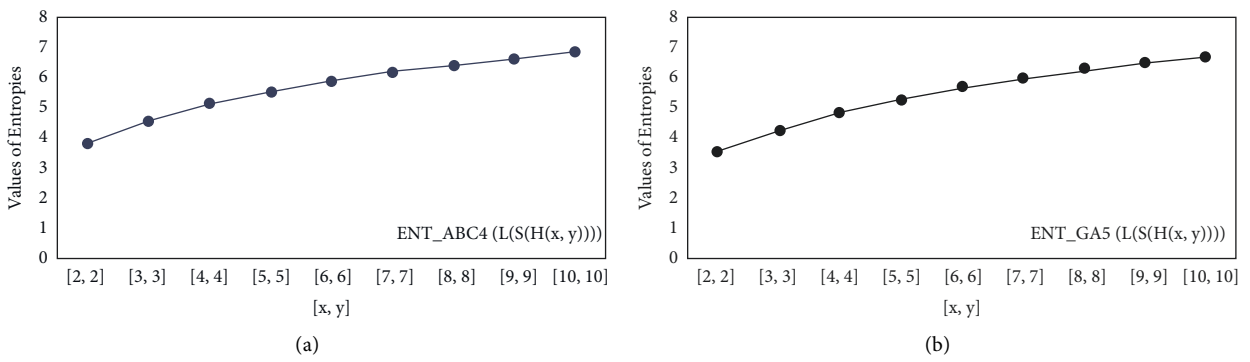


FIGURE 12: (a) The ABC_4 entropy, (b) The GA_5 entropy $x = 1, y \neq 1$.

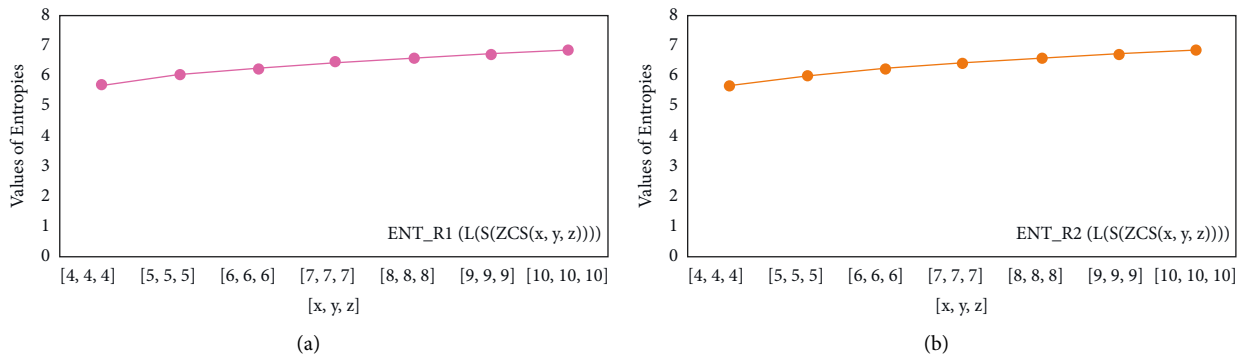


FIGURE 13: (a) R_1 entropy, (b) R_{-1} entropy.

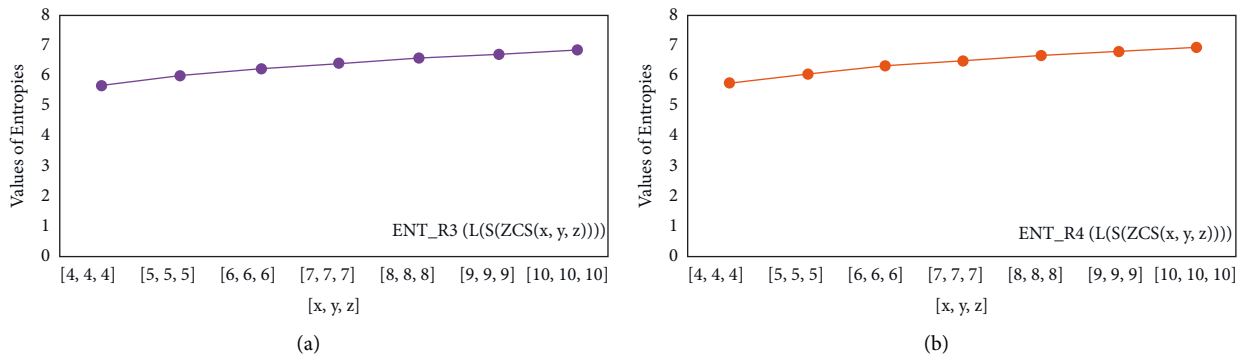


FIGURE 14: (a) $R_{1/2}$ entropy, (b) $R_{1/2}$ entropy.

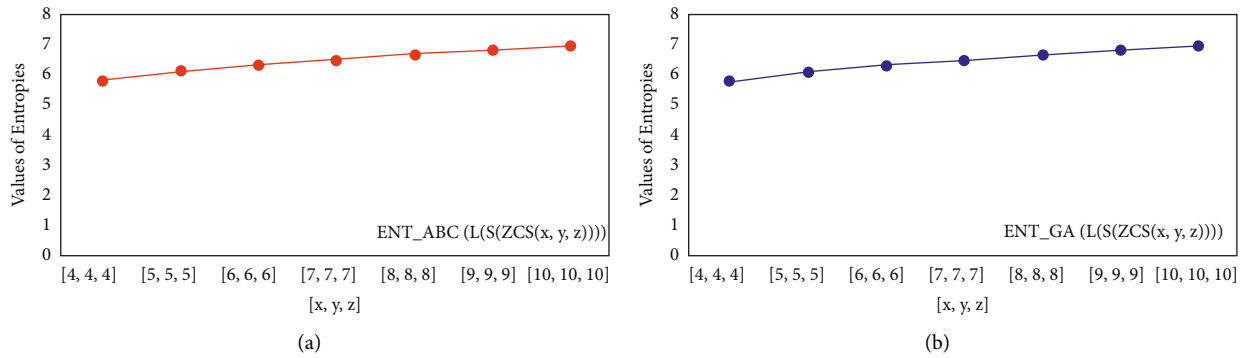
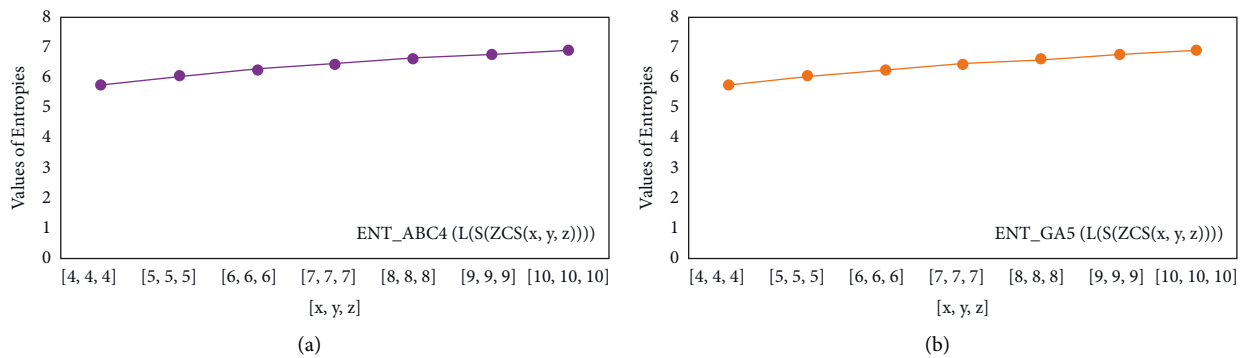


FIGURE 15: (a)ABC entropy, (b)The GA, entropy.


 FIGURE 16: (a)ABC₄ entropy, (b)GA₅ entropy.

4.1.5. *The GA₅ Entropy of \mathcal{F} .* After some simple calculations, the GA₅ index and corresponding entropy measure with the help of Table 7 and equation (11) subject to the condition that $x = y = z \geq 4$.

$$\begin{aligned}
 GA_5(\mathcal{F}) &= (x + y + z) \left(15 + \frac{48\sqrt{10}}{13} + \frac{144\sqrt{2}}{17} \right) + 3 + \frac{16\sqrt{5}}{3} - \frac{336\sqrt{10}}{13} - \frac{720\sqrt{2}}{17}, \\
 ENT_{GA_5}(\mathcal{F}) &= \log(GA_5) - \frac{1}{(GA_5)} \sum_{i=1}^7 \sum_{lm \in E_i(\mathcal{F})} \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right] \log \left[\sqrt{\frac{A_l + A_m - 2}{A_l A_m}} \right], \\
 ENT_{GA_5}(\mathcal{F}) &= \log(GA_5) - \frac{[16\sqrt{5}/3] \log [4\sqrt{5}/9]}{(GA_5)} - \frac{[48\sqrt{10}/13] (x + y + z - 7) \log [4\sqrt{10}/13]}{(GA_5)} \\
 &\quad - \frac{[144\sqrt{2}/17] (x + y + z - 5) \log [12\sqrt{2}/17]}{(GA_5)}.
 \end{aligned} \tag{38}$$

5. Concluding remarks for Computed Results

The applications of information-theoretic framework in many disciplines of study, such as biology, physics, engineering, and social sciences, have grown exponentially in the recent two decades. This phenomenal increase has been particularly impressive in the fields of soft computing, molecular biology, and information technology. As a result, the scientists may find our numerical and

graphical results useful [54, 55]. The entropy function is monotonic, which means that as the size of a chemical structure increases, so does the entropy measure, and as the entropy of a system increases, so does the uncertainty regarding its reaction.

For $L(S(T_x))$, the numerical and graphical results are shown in Tables 8 and 9 and Figures 4–7. In Table 9, the fifth arithmetic geometric entropy is zero which shows that the process is deterministic for $x = 1$. When the chemical

structure $L(S(T_x))$ expands, the Randić entropy for $\alpha = 1/2$ develops more quickly than other entropy measurements of $L(S(T_x))$, whereas the Randić entropy for $\alpha = -1/2$ develops more slowly. This demonstrates that different topologies have varied entropy characteristics. For $L(S(H(x, y)))$, the numerical and graphical results are shown in Tables 10–13 and Figures 8–12. When the chemical structure $L(S(H(x, y)))$ expands, the geometric arithmetic entropy develops more quickly than other entropy measurements of $L(S(H(x, y)))$, whereas the ABC_4 entropy develops more slowly. Finally, for $L(S(ZCS(x, y, z)))$, the numerical and graphical results are shown in Table 14 and Figures 13–16. When the chemical structure $L(S(ZCS(x, y, z)))$ expands, the geometric arithmetic entropy develops more quickly than other entropy measurements of $L(S(ZCS(x, y, z)))$, whereas the Randić entropy for $\alpha = -1$ develops more slowly.

The novelty of this article is that entropies are computed for three types of benzenoid systems. These entropy measures are useful in estimating the heat of formation and many Physico-chemical properties. In statistical analysis of benzene structures, entropy measures showed more significant results as compared to topological indices. Therefore, we can say that the entropy measure is a newly introduced topological descriptor.

6. Conclusion

Using Shannon's entropy and Chen et al. [31] entropy definitions, we generated graph entropies associated to a new information function in this research. Between indices and information entropies, a relationship is created. Using the line graph of the subdivision of these graphs, we estimated the entropies for triangular benzenoids T_x , hexagonal parallelogram $H(x, y)$ nanotubes, and $ZCS(x, y, z)$. Thermodynamic entropy of enzyme-substrate complexions [57, 58] and configuration entropy of glass-forming liquids [56] are two examples of thermodynamic entropy employed in molecular dynamics studies of complex chemical systems. Similarly, using information entropy as a crucial structural criterion could be a new step in this direction.

Data Availability

The data used to support the findings of this study are cited at relevant places within the text as references.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

This work was equally contributed by all writers.

References



- [1] Z. Q. Chu, M. K. Siddiqui, S. Manzoor, S. A. K. Kirmani, M. F. Hanif, and M. H. Muhammad, "On rational curve fitting between topological indices and entropy measures for graphite carbon nitride," *Polycyclic Aromatic Compounds*, vol. 5, pp. 1–18, 2022.
- [2] W. Gao and M. R. Farahani, "Degree-based indices computation for special chemical molecular structures using edge dividing method," *Applied Mathematics and Nonlinear Sciences*, vol. 1, no. 1, pp. 99–122, 2016.
- [3] E. Estrada, L. Torres, L. Rodríguez, and I. Gutman, "An atom-bond connectivity index. Modelling the enthalpy of formation of alkanes," *Indian Journal of Chemistry*, vol. 37A, pp. 849–855, 1998.
- [4] E. Estrada and N. Hatano, "Statistical-mechanical approach to subgraph centrality in complex networks," *Chemical Physics Letters*, vol. 439, no. 1-3, pp. 247–251, 2007.
- [5] E. Estrada, "Generalized walks-based centrality measures for complex biological networks," *Journal of Theoretical Biology*, vol. 263, no. 4, pp. 556–565, 2010.
- [6] W. Gao, M. K. Siddiqui, M. Naeem, and N. A. Rehman, "Topological characterization of carbon graphite and crystal cubic carbon structures," *Molecules*, vol. 22, no. 9, pp. 1496–1507, 2017.
- [7] M. Imran, M. K. Siddiqui, M. Naeem, and M. A. Iqbal, "On topological properties of symmetric chemical structures," *Symmetry*, vol. 10, no. 5, pp. 173–221, 2018.
- [8] M. Randić, "On characterization of molecular branching," *Journal of the American Chemical Society*, vol. 97, no. 23, pp. 6609–6615, 1975.
- [9] M. K. Siddiqui, M. Imran, and A. Ahmad, "On Zagreb indices, Zagreb polynomials of some nanostar dendrimers," *Applied Mathematics and Computation*, vol. 280, pp. 132–139, 2016.
- [10] X. Zhang, X. Wu, S. Akhter, M. K. Jamil, J. B. Liu, and M. R. Farahani, "Edge-version atom-bond connectivity and geometric arithmetic indices of generalized bridge molecular graphs," *Symmetry*, vol. 10, no. 12, pp. 751–786, 2018.
- [11] X. Zhang, H. M. Awais, M. Javaid, and M. K. Siddiqui, "Multiplicative Zagreb indices of molecular graphs," *Journal of Chemistry*, vol. 2019, Article ID 5294198, 19 pages, 2019.
- [12] X. Zhang, A. Rauf, M. Ishtiaq, M. K. Siddiqui, and M. H. Muhammad, "On degree based topological properties of two carbon nanotubes," *Polycyclic Aromatic Compounds*, vol. 42, no. 3, pp. 866–884, 2020.
- [13] X. Zhang, H. Jiang, J. B. Liu, and Z. Shao, "The cartesian product and join graphs on edge-version atom-bond connectivity and geometric arithmetic indices," *Molecules*, vol. 23, no. 7, pp. 1731–1817, 2018.
- [14] X. Zhang, M. Naeem, A. Q. Baig, and M. A. Zahid, "Study of hardness of superhard crystals by topological indices," *Journal of Chemistry*, vol. 2021, Article ID 9604106, 10 pages, 2021.
- [15] X. Zhang, M. K. Siddiqui, S. Javed, L. Sherin, F. Kausar, and M. H. Muhammad, "Physical analysis of heat for formation and entropy of Ceria Oxide using topological indices," *Combinatorial Chemistry & High Throughput Screening*, vol. 25, no. 3, pp. 441–450, 2022.
- [16] M. K. Siddiqui, M. Naeem, N. A. Rahman, and M. Imran, "Computing topological indices of certain networks," *Journal of Optoelectronics and Advanced Materials*, vol. 18, no. 9-10, pp. 884–892, 2016.
- [17] D. Vukicevic and B. Furtula, "Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges," *Journal of Mathematical Chemistry*, vol. 46, no. 4, pp. 1369–1376, 2009.
- [18] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

- [19] E. Trucco, "A note on the information content of graphs," *Bulletin of Mathematical Biophysics*, vol. 18, no. 2, pp. 129–135, 1956.
- [20] A. Mowshowitz, "Entropy and the complexity of graphs: I. An index of the relative complexity of a graph," *Bulletin of Mathematical Biophysics*, vol. 30, no. 1, pp. 175–204, 1968.
- [21] N. Rashevsky, "Life, information theory, and topology," *Bulletin of Mathematical Biophysics*, vol. 17, no. 3, pp. 229–235, 1955.
- [22] G. Karreman, "Topological information content and chemical reactions," *Bulletin of Mathematical Biophysics*, vol. 17, no. 4, pp. 279–285, 1955.
- [23] D. Bonchev, *Complexity in Chemistry, Introduction and Fundamentals*, vol. 2, pp. 1–100, Taylor & Francis, Boca Raton, FL, USA, 2003.
- [24] D. G. Bonchev, "Kolmogorov's information, shannon's entropy, and topological complexity of molecules," *Bulgarian Chemical Communications*, vol. 28, no. 3-4, pp. 567–582, 1995.
- [25] D. Bonchev and N. Trinajstić, "Information theory, distance matrix, and molecular branching," *The Journal of Chemical Physics*, vol. 67, no. 10, pp. 4517–4533, 1977.
- [26] D. Bonchev, O. V. Mekenyan, and N. Trinajstić, "Isomer discrimination by topological information approach," *Journal of Computational Chemistry*, vol. 2, no. 2, pp. 127–148, 1981.
- [27] G. Castellano and F. Torrens, "Information entropy-based classification of triterpenoids and steroids from Ganoderma," *Phytochemistry*, vol. 116, pp. 305–313, 2015.
- [28] M. Dehmer, K. Varmuza, S. Borgert, and F. Emmert-Streib, "On entropy-based molecular descriptors: statistical analysis of real and synthetic chemical structures," *Journal of Chemical Information and Modeling*, vol. 49, no. 7, pp. 1655–1663, 2009.
- [29] Z. Terenteva and N. I. Kobozev, "Possible relation between entropy of information-theory and thermodynamic entropy," *Zhurnal Fizicheskoi Khimii*, vol. 50, no. 4, pp. 877–881, 1976.
- [30] Y. A. Zhdanov, *Information Entropy in Organic Chemistry*, vol. 1, Rostov University, Rostov Oblast, Russia, 1979, in Russian.
- [31] Z. Chen, M. Dehmer, and Y. Shi, "A note on distance based graph entropies," *Entropy*, vol. 16, no. 10, pp. 5416–5427, 2014.
- [32] S. Cao, M. Dehmer, and Y. Shi, "Extremality of degree-based graph entropies," *Information Sciences*, vol. 278, pp. 22–33, 2014.
- [33] M. Dehmer, "Information processing in complex networks: graph entropy and information functionals," *Applied Mathematics and Computation*, vol. 201, no. 1-2, pp. 82–94, 2008.
- [34] M. Dehmer and A. Mowshowitz, "A history of graph entropy measures," *Information Sciences*, vol. 181, no. 1, pp. 57–78, 2011.
- [35] S. Manzoor, M. K. Siddiqui, and S. Ahmad, "On entropy measures of molecular graphs using topological indices," *Arabian Journal of Chemistry*, vol. 13, no. 8, pp. 6285–6298, 2020.
- [36] S. Cao and M. Dehmer, "Degree-based entropies of networks revisited," *Applied Mathematics and Computation*, vol. 261, pp. 141–147, 2015.
- [37] M. Dehmer, L. Sivakumar, and K. Varmuza, "Uniquely discriminating molecular structures using novel eigenvalue-based descriptors," *MATCH Communications in Mathematical*, vol. 67, pp. 147–172, 2012.
- [38] H. J. Morowitz, "Some order-disorder considerations in living systems," *Bulletin of Mathematical Biophysics*, vol. 17, no. 2, pp. 81–86, 1955.
- [39] N. Rashevsky, "Information theory in biology," *Bulletin of Mathematical Biophysics*, vol. 16, no. 2, pp. 183–185, 1954.
- [40] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE - Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.
- [41] R. V. Sol and S. I. Valverde, "Information theory of complex networks: on evolution and architectural constraints," *Complex Networks Lectures Notes in Physics*, vol. 650, pp. 189–207, 2004.
- [42] Y. J. Tan and J. Wu, "Network structure entropy and its application to scale-free networks," *Systems Engineering - Theory & Practice*, vol. 6, pp. 1–3, 2004.
- [43] S. Manzoor, M. K. Siddiqui, and S. Ahmad, "On physical analysis of degree-based entropy measures for metal-organic superlattices," *The European Physical Journal Plus*, vol. 136, no. 3, pp. 287–322, 2021.
- [44] S. Manzoor, M. K. Siddiqui, and S. Ahmad, "Degree-based entropy of molecular structure of hyaluronic acid-curcumin conjugates," *The European Physical Journal Plus*, vol. 136, no. 1, pp. 15–21, 2021.
- [45] M. A. Rashid, S. Ahmad, M. K. Siddiqui, S. Manzoor, and M. Dhlamini, "An Analysis of Eccentricity-Based Invariants for Biochemical Hypernetworks," *Complexity*, vol. 2021, Article ID 1974642, 14 pages, 2021.
- [46] S. Akhter and M. Imran, "On molecular topological properties of benzenoid structures," *Canadian Journal of Chemistry*, vol. 94, no. 8, pp. 687–698, 2016.
- [47] N. K. Raut, "Topological indices and topological polynomials of triangular benzenoid system," *International Journal of Mathematics Trends and Technology*, vol. 67, no. 6, pp. 90–96, 2021.
- [48] Z. Hussain, S. Rafique, M. Munir et al., "Irregularity molecular descriptors of hourglass, jagged-rectangle, and triangular benzenoid systems," *Processes*, vol. 7, no. 7, p. 413, 2019.
- [49] Y. C. Kwun, A. Ali, W. Nazeer, M. Ahmad Chaudhary, and S. M. Kang, "M-polynomials and degree-based topological indices of triangular, hourglass, and jagged-rectangle benzenoid systems," *Journal of Chemistry*, vol. 2018, Article ID 8213950, 8 pages, 2018.
- [50] V. R. Kulli, B. Chaluvaraju, and H. S. Boregowda, "Connectivity Bhatti indices for certain families of benzenoid systems," *Journal of Ultra Chemistry*, vol. 13, no. 04, pp. 81–87, 2017.
- [51] M. Ghorbani and M. Ghazi, "Computing some topological indices of Triangular Benzenoid," *Digest Journal of Nanomaterials and Biostructures*, vol. 5, no. 4, pp. 1107–1111, 2010.
- [52] A. Q. Baig, M. Imran, and H. Ali, "Computing Omega, Sadhana and PI polynomials of benzenoid carbon nanotubes," *Optoelectronics and advanced materials-rapid communications*, vol. 9, pp. 248–255, 2015.
- [53] K. G. Mirajkar and B. Pooja, "On Gourava indices of some chemical graphs," *International Journal of Applied Engineering Research*, vol. 14, no. 3, pp. 743–749, 2019.
- [54] A. Mehler, A. Lücking, and P. Weib, "A network model of interpersonal alignment in dialog," *Entropy*, vol. 12, no. 6, pp. 1440–1483, 2010.
- [55] A. Mowshowitz and M. Dehmer, "Entropy and the complexity of graphs revisited," *Entropy*, vol. 14, no. 3, pp. 559–570, 2012.
- [56] Y. Champion and N. Thurieau, "The sample size effect in metallic glass deformation," *Scientific Reports*, vol. 10, no. 1, Article ID 10801, 2020.
- [57] M. M. H. Graf, U. Bren, D. Haltrich, and C. Oostenbrink, "Molecular dynamics simulations give insight into d-glucose dioxidation at C2 and C3 by *Agaricus meleagris* pyranose

- dehydrogenase C-2 and C-3 by *Agaricus meleagris* pyranose dehydrogenase," *Journal of Computer-Aided Molecular Design*, vol. 27, no. 4, pp. 295–304, 2013.
- [58] M. V. Putz, A. M. Lacrama, and V. Ostafe, "Full analytic progress curves of enzymic reactions in vitro," *International Journal of Molecular Sciences*, vol. 7, no. 11, pp. 469–484, 2006.
- [59] M. Dehmer and M. Graber, "The discrimination power of molecular identification numbers revisited," *MATCH Communications in Mathematical*, vol. 69, pp. 785–794, 2013.
- [60] R. E. Ulanowicz, "Quantitative methods for ecological network analysis," *Computational Biology and Chemistry*, vol. 28, no. 5-6, pp. 321–339, 2004.

Research Article

Generalized β – Hyers–Ulam–Rassias Stability of Impulsive Difference Equations

Yahya Almalki ¹, Gul Rahmat,² Atta Ullah,² Fatima Shehryar,² Muhammad Numan,³ and Muhammad Usman Ali ³

¹Department of Mathematics, College of Sciences, King Khalid University, Abha 61413, Saudi Arabia

²Department of Mathematics, Islamia College Peshawar, Peshawar, Pakistan

³Department of Mathematics, COMSATS University Islamabad, Attock Campus, Attock, Pakistan

Correspondence should be addressed to Muhammad Usman Ali; muh_usman_ali@yahoo.com

Received 23 May 2022; Revised 6 July 2022; Accepted 30 July 2022; Published 15 September 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yahya Almalki 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 describes the existence and uniqueness of the solution, β -Hyers–Ulam–Rassias stability and generalized β -Hyers–Ulam–Rassias stability of an impulsive difference system on bounded and unbounded discrete intervals. At the end, an example is given to illustrate the theoretical result.

1. Introduction

Many physical problems can be expressed in mathematical models using differential equations. Differential equations enable us to study the rapid changes in physical problems, for example, blood flows, river flows, biological systems, control theory, and mechanical systems with impact. A system of differential equations with impulses can be used to model several above-listed problems. A few existing results for a general class of impulsive systems were discussed by Ahmad [1]. The theory of impulsive difference equations was studied in [2–4]. In [5], the existence of solutions for semilinear abstract differential equations without instantaneous impulses was discussed.

At the University of Wisconsin, Ulam [6] proposed the stability problem, stated as follows. Let us denote by H_1 the group and by H_2 the metric group with a metric δ and a constant $\nu > 0$. The problem is to study if there exists $\lambda > 0$ satisfies for every $h: H_1 \rightarrow H_2$ such that

$$\delta(h(\sigma\nu), h(\sigma)h(\nu)) \leq \lambda, \quad \forall \sigma, \nu \in H_1, \quad (1)$$

there exists a homomorphism $f: H_1 \rightarrow H_2$ that satisfies

$$\delta(h(\sigma), f(\sigma)) \leq \nu, \quad \forall \sigma \in H_1. \quad (2)$$

The linear functional equations, of the form $f(x + y) = f(x) + f(y)$, and their solutions have been discussed in several spaces. A linear transformation is a solution of a linear functional equation. By considering the H_1 and H_2 as Banach spaces, Hyers [7] discussed the above problem in terms of linear functional equations. Then, Aoki [8] and Rassias [9] extended the concept of Hyers and Ulam. In the last decade, we have seen some worthwhile generalizations in the direction of Ulam stability.

In 2012, Wang et al. [10] studied the Ulam-type stability of first-order nonlinear impulsive differential equations by utilizing the bounded interval with finite impulses. In 2014, Wang et al. [11] studied the Hyers–Ulam–Rassias stability and generalized Hyers–Ulam–Rassias stability for impulsive evolution equations on a closed and bounded interval. In 2015, Zada et al. [12] studied the Hyers–Ulam stability of differential systems in terms of a dichotomy. The existence and Hyers–Ulam stability of the periodic fractional stochastic and Riemann–Liouville fractional neutral functional stochastic impulsive differential equations were given [13, 14]. Recently, Rahmat et al. [15] studied the Hyers–Ulam stability of delay differential equations. In 2019, Hu and Zhu [16] presented the stability criteria for an impulsive stochastic functional differential system with

distributed delay-dependent impulsive effects. Furthermore, Hu et al. [17] provided the improved Razumikhin stability criteria for an impulsive stochastic delay differential system, and for a detail study, we refer to the readers to [17] and the references therein.

In this paper, we will explain the β -Hyers–Ulam–Rassias stability and generalized β -Hyers–Ulam–Rassias stability of the impulsive difference system of the form

$$\begin{cases} \Theta_{n+1} = H\Theta_n + B\zeta_n + f(n, \Theta_n, \zeta_n), & n \geq 0, \\ \Theta_0, & n = 0, \\ \Theta_{n_k+1} = \Theta_{n_k-1} + I_k(n, \Theta_{n_k-1}, \zeta_{n_k-1}), & k = 1, 2, 3, \dots, m, \end{cases} \quad (3)$$

where the constant matrix $H, B \in \mathbb{R}^{n \times n}$, $f \in \mathbb{C}(\mathbb{Z}_{+ \times X, X})$ and $\Theta_n \in B(\mathbb{Z}_+, \mathbb{X})$ space of bounded and convergent sequences, $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ and $\mathbb{X} = \mathbb{R}^n$, $I = \{0, 1, 2, \dots, n\}$. In fact, we are presenting a discrete version of the work given in [18], in which β -Hyers–Ulam–Rassias stability was discussed for differential equations. With the help of [15, 18], we find out β -Hyers–Ulam–Rassias stability of the difference equation.

2. Preliminaries

Here, we discuss some notation and definitions, which will be needed for our main work. The n -dimensional Euclidean space will be denoted by \mathbb{R}^n along with the vector norm $\|\cdot\|$, and $n \times n$ matrices with real-valued entries will be denoted by $\mathbb{R}^{n \times n}$. The vector infinite norm is defined as $\|v\| = \max_{1 \leq i \leq n} |v_i|$, and the matrix infinite-norm is given as $\|A\| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ where $v \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$, also v_i and a_{ij} are the elements of the vector v and the matrix A , respectively. $\mathbb{C}(\mathbf{I}, \mathbb{X})$ will be the space of all convergent sequences from \mathbf{I} to \mathbb{X} with norm $\|v\| = \sup_{n \in \mathbf{I}} \|v_n\|$. We will use \mathbb{R}, \mathbb{Z} , and \mathbb{Z}_+ for the set of all real, integer, and non-negative integer numbers, respectively. The next lemma is a basic result about the solution of the difference system (1).

Lemma 1. *The impulsive difference system (1) has the solution*

$$\begin{aligned} X_n = & H^n \Theta_0 + H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \\ & + \sum_{n_k=0}^n T(n - n_k) I_k(n_k, \Theta_{n_k}, \zeta_{n_k}), \quad n \in I. \end{aligned} \quad (4)$$

The solution can easily be obtained by consecutively placing the values of $n \in \{0, 1, 2, \dots\}$.

Definition 1. A function $\|\cdot\|_\beta: \mathbb{V} \rightarrow [0, \infty)$ is called β -norm, with $0 < \beta \leq 1$, where \mathbb{V} is a vector space over the field \mathbf{K} , if the function satisfied the following properties:

- (1) $\|\mathcal{H}\|_\beta = 0$ if and only $\mathcal{H} = 0$
- (2) $\|\kappa \mathcal{H}\|_\beta = |\kappa|^\beta \|\mathcal{H}\|_\beta$, for each $\kappa \in \mathbf{K}$ and $\mathcal{H} \in \mathbb{V}$
- (3) $\|\mathcal{H} + \mathcal{H}_1\|_\beta \leq \|\mathcal{H}\|_\beta + \|\mathcal{H}_1\|_\beta$, for all $\mathcal{H}, \mathcal{H}_1 \in \mathbb{V}$

And $(\mathbb{V}, \|\cdot\|_\beta)$ is said to be β -norm space.

Definition 2. Let $\epsilon > 0, \psi > 0$ and $\varphi_n \in B(I, X)$. A sequence Θ_n will be an ϵ -approximate solution of (1), if

$$\begin{cases} \|\Theta_{n+1} - H\Theta_n - B\zeta_n - f(n, \Theta_n, \zeta_n)\| \leq \epsilon \varphi_n, & n \geq 0, \\ \|\Theta_{n_k+1} - \Theta_{n_k-1} - I_k(n, \Theta_{n_k-1}, \zeta_{n_k-1})\| \leq \epsilon \psi, & k = 1, 2, 3, \dots, m, \end{cases} \quad (5)$$

Definition 3. System (1) is said to be β -Hyers–Ulam–Rassias stable if for every ϵ -approximate solution Y_n of system (1), there exists an exact solution Θ_n of (1) and a nonnegative real number $\mathcal{C}_{f, N, M, \varphi, \Psi}$ such that

$$\|Y_n - \Theta_n\|^\beta \leq \mathcal{C}_{f, N, M, \varphi, \Psi} \epsilon^\beta (\varphi_n^\beta + \psi^\beta), \text{ for all } n \in I. \quad (6)$$

Definition 4. System (1) is a generalized Hyers–Ulam–Rassias stable if for every ϵ -approximate solution Y_n of system (1), there will be an exact solution Θ_n of (1) and a nonnegative real scalar $\mathcal{F}_{M, \eta, \varphi, \eta_\varphi, f}$ such that

$$\|Y_n - \Theta_n\|^\beta \leq \mathcal{F}_{M, \eta, \varphi, \eta_\varphi, f} \epsilon^\beta (\varphi_n^\beta + \Psi_{k+1}^\beta), \quad n \in I. \quad (7)$$

Remark 1. From (2), it is clear that $Y \in \mathbb{C}(\mathbf{I}, \mathbf{X})$ satisfies (2) if and only if there exists $h \in \mathbb{C}(\mathbf{I}, \mathbf{X})$ and a sequence $h_k, k \in M$ satisfying

$$\begin{cases} \|h_n\| \leq \epsilon \psi_n, & n \in M, \\ \Theta_{n+1} = H\Theta_n + B\zeta_n + \mathbf{f}(n, \Theta_n, \zeta_n) + h_n, & n \in \mathbb{Z}_+, \\ Y_0 = \Theta_0 + h_0, \\ \Theta_{n_k+1} - \Theta_{n_k-1} = I_k(\Theta_{n_k-1}, \zeta_{n_k-1}) + h_{n_k}, & k \in M. \end{cases} \quad (8)$$

The solution of Remark 1 is

$$\begin{aligned} \Theta_n = & H^n (\Theta_0 + h_0) + H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i \\ & + H^{n-1} \sum_{i=0}^{n-1} H^{-i} (f(i, \Theta_i, \zeta_i) + h_i) \\ & + \sum_{n_k=0}^n T(n - n_k) (I_k(n_k, \Theta_{n_k}, \zeta_{n_k}) + h_{n_k}), \quad n \in I. \end{aligned} \quad (9)$$

Lemma 2 (see [19]). *for any $n \geq 0$ with*

$$\|U_n\| \leq a_n + \sum_{i=0}^n P_i U_i + \sum_{0 \leq n_k \leq n} \gamma_k U_{n_k-1}, \quad n \geq 0, \quad (10)$$

then, we have

$$U_n \leq a_n (1 + \gamma_k)^k \exp\left(\sum_{i=0}^n P_i\right), \quad \text{where } k \in I. \quad (11)$$

Remark 2. If we replace γ_k by γ_{k_n} , then

$$Y_n \leq c \prod_{n_k=0}^n (1 + \gamma_{k_n}) \exp\left(\sum_{i=0}^n P_i\right) \text{ for } n \geq 0. \quad (12)$$

G_2^* :

$$\left[M \sum_{i=0}^{n-1} \|H^{-i}\| \mathcal{L}_f + Nn\mathcal{L}_{I_k} \right] < 1. \quad (15)$$

3. Uniqueness and Existence of Solution of an Impulsive Difference System

To describe the uniqueness and existence of the solution of system (1), we will use the following assumptions:

G_1 : for $f, I_k: \mathbb{Z}_{+ \times X \times X} \rightarrow \mathbb{X}$, $k \in I$, there exist constants $\mathcal{L}_f > 0$ and $\mathcal{L}_{I_k} > 0$, such that

$$\begin{aligned} \|f(n, \mathfrak{F}, p) - f(n, \mathfrak{F}', p)\| &\leq \mathcal{L}_f \|\mathfrak{F} - \mathfrak{F}'\|, \\ \|I_k(n, w, y) - I_k(n, w', y)\| &\leq \mathcal{L}_{I_k} \|w - w'\|. \end{aligned} \quad (13)$$

G_2 :

$$M = \sup_{i \in I} \|H^i\|, \quad N = \max_{0 \leq n_k \leq n} \|T(n - n_k)\|. \quad (14)$$

Theorem 1. *If assumptions G_1 , G_2 , and G_2^* are held, then system (1) has a unique solution $\Theta \in \mathbb{C}(I, \mathbb{X})$.*

Proof. Define $\mathcal{A}: \mathbb{C}(I, \mathbb{X}) \rightarrow \mathbb{C}(I, \mathbb{X})$ by

$$\begin{aligned} \mathcal{A}\Theta_n &= H^n \Theta_0 + H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \\ &\quad + \sum_{n_k=0}^n T(n - n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}). \end{aligned} \quad (16)$$

Now, for $\Theta, \Theta' \in \mathbb{C}(I, \mathbb{X})$, we have

$$\begin{aligned} \|\mathcal{A}\Theta_n - \mathcal{A}\Theta'_n\| &= \left\| \begin{aligned} &H^n \Theta_0 + H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \\ &\quad + \sum_{n_k=0}^n T(n - n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) \\ &- H^n \Theta'_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta'_i, \zeta_i) \\ &\quad - \sum_{n_k=0}^n T(n - n_k) I_k(n_k, \Theta'_{n_k-1}, \zeta_{n_k-1}) \end{aligned} \right\|. \end{aligned} \quad (17)$$

This implies that

$$\begin{aligned} \|\mathcal{A}\Theta_n - \mathcal{A}\Theta'_n\| &= \left\| \begin{aligned} &H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) + \sum_{n_k=0}^n T(n - n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) \\ &- H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta'_i, \zeta_i) - \sum_{n_k=0}^n T(n - n_k) I_k(n_k, \Theta'_{n_k-1}, \zeta_{n_k-1}) \end{aligned} \right\| \\ &\leq + \sum_{n_k=0}^n \|T(n - n_k)\| \left\| I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) - I_k(n_k, \Theta'_{n_k-1}, \zeta_{n_k-1}) \right\| \\ &\leq M \sum_{i=0}^{n-1} \|H^{-i}\| \mathcal{L}_f \|\Theta_i - \Theta'_i\| + N \sum_{n_k=0}^n \mathcal{L}_{I_k} \|\Theta_{n_k-1} - \Theta'_{n_k-1}\| \\ &\leq \left[M \sum_{i=0}^{n-1} \|H^{-i}\| \mathcal{L}_f + Nn\mathcal{L}_{I_k} \right] \|\Theta - \Theta'\|. \end{aligned} \quad (18)$$

This implies that \mathcal{A} is a contraction map using the Banach contraction principle, we say that system (1) has a unique solution. \square

4. β -Hyers–Ulam–Rassias Stability on Bounded Discrete Interval

To determine β -Hyers–Ulam–Rassias stability on the bounded discrete interval, we have one more assumption:

G_3 : there exist a constant $\eta_\phi > 0$ and φ_n and a nondecreasing function $\varphi \in B(I, X)$ such that

$$\sum_{i=0}^{n-1} \psi_i \leq \eta_\phi \phi_n. \quad (19)$$

Theorem 2. *System (1) is β -Hyers–Ulam–Rassias stable over discrete bounded interval, if G_1 , G_2 and G_3 are satisfied.*

Proof. The solution of system (1) is as follows:

$$\begin{aligned} \Theta_n &= H^n \Theta_0 + H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \\ &\quad + \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}). \end{aligned} \quad (20)$$

Let Y_n be the solution of inequality (2), we have

$$\begin{aligned} &\left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^n H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^n H^{-i} f(i, Y_i, \zeta_i) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right\| \\ &= \left\| H^n h_n + H^{n-1} \sum_{i=0}^{n-1} H^{-i} h_i + \sum_{n_k=0}^n T(n-n_k) h_{n_k} \right\| \leq M \varepsilon \Psi + M^2 \delta \sum_{i=0}^{n-1} \Psi_i + N \varepsilon \sum_{n_k=0}^m \Psi \\ &= (M+N) \sum_{k=0}^m \delta \Psi + \sum_{i=0}^{n-1} M^2 \delta \Psi_i \\ &\leq (mM\delta + M^2 \varepsilon \eta_\phi) (\varphi_n + \varphi) \\ &= M\delta (m + M\eta_\phi) (\varphi_n + \varphi), \end{aligned} \quad (21)$$

thus, for each $n \in \{n_k, n_{k+1}, \dots\}$, we have

$$\begin{aligned} \|Y_n - \theta_n\|^\beta &= \left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \right. \\ &\quad \left. - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) \right\|^\beta \\ &= \left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \right. \\ &\quad \left. - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) - H^{n-1} \right. \\ &\quad \left. \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) + \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right. \\ &\quad \left. - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right\|^\beta \end{aligned}$$

$$\begin{aligned}
& \leq \left(\left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right\| \right)^\beta \\
& \quad + \left(\left\| H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \right\| \right)^\beta \\
& \quad + \left(\left\| \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) \right\| \right)^\beta \\
& \leq (M\varepsilon(m + M\eta_\phi)(\varphi_n + \varphi))^\beta + \left(M^2 \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \|Y_i - \Theta_i\| \right)^\beta + \left(N \sum_{k=1}^m \mathcal{L}_{I_{n_k}} \|Y_{n_k} - \Theta_{n_k}\| \right)^\beta \\
& \quad \|Y_n - \Theta_n\| \leq 3^{(1/\beta)-1} \left[M\varepsilon(m + M\eta_\phi)(\varphi_n + \varphi) + M^2 \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \|Y_i - \Theta_i\| + N \sum_{k=1}^m \mathcal{L}_{I_{n_k}} \|Y_{n_k} - \Theta_{n_k}\| \right], \tag{22}
\end{aligned}$$

by using the relation

$$\begin{aligned}
(x + y + z)^\gamma & \leq 3^{\gamma-1} (x^\gamma + y^\gamma + z^\gamma), \\
\text{where } k \quad x, y, z & \geq 0 \quad \text{and } \gamma > 1. \tag{23}
\end{aligned}$$

Now, using Gronwall Lemma 2, we get

$$\|Y_N - \Theta_n\| \leq 3^{(1/\beta)-1} [M\varepsilon(m + M\eta_\phi)(\varphi_n + \varphi)] (1 + 3^{(1/\beta)-1} N \mathcal{L}_{I_k})^k \exp \left(M^2 3^{(1/\beta)-1} \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \right). \tag{24}$$

Now,

$$\begin{aligned}
\|Y_n - \Theta_n\|^\beta & \leq 3^{1-\beta} [M\varepsilon(m + M\eta_\phi)(\varphi_n + \varphi)]^\beta (1 + 3^{(1/\beta)-1} N \mathcal{L}_{I_k})^{k\beta} \exp \left(M^2 \beta 3^{(1/\beta)-1} \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \right) \\
& \leq 3^{1-\beta} (M\varepsilon(m + M\eta_\phi))^\beta (\varphi_n + \varphi)^\beta (1 + 3^{(1/\beta)-1} N \mathcal{L}_{I_k})^{k\beta} \exp \left(M^2 \beta 3^{(1/\beta)-1} \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \right) \\
& \leq C_{f,N,M,\varphi,\Psi} \varepsilon^\beta ((\varphi_n)^\beta + \varphi^\beta), \tag{25}
\end{aligned}$$

using $(x + y)^r \leq (x^r + y^r)$, $x, y \geq 0$, for any $r \in \{0, 1\}$, where

$$C_{f,N,M,\varphi,\Psi} = 3^{1-\beta} (M\varepsilon(m + M\eta_\phi))^\beta (1 + 3^{(1/\beta)-1} N \mathcal{L}_{I_k})^{k\beta} \exp \left(M^2 \beta 3^{(1/\beta)-1} \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \right). \tag{26}$$

Hence, system (1) is β -Hyers–Ulam–Rassias stable. \square

5. β -Hyers–Ulam–Rassias Stability on Unbounded Discrete Interval

To explain β -Hyers–Ulam–Rassias stability on an unbounded discrete interval, we must need the following assumptions:

G_4 : the operators family $\|H^{n-i}\| \leq Me^{\omega(n-i)}$ and $\|T(n-n_k)\| \leq Me^{\omega(n-n_k)}$.

$$\begin{aligned} G_5: \sum_{i=0}^{n-1} \mathcal{L}_{f_i} &\leq \kappa_f n + \kappa_f, \\ G_6: \mathcal{L}_{\mathcal{J}} &= \left(1 + 3^{(1/\beta)-1} M \mathcal{L}_{I_{n_k}}\right)^k < \infty, \\ G_7: \sum_{i=0}^{n-1} e^{\omega(n-i)+3^{1/\beta-1} M \kappa_f n} \Psi_i &\leq \eta_\varphi \varphi_n, \\ G_8: M_1 &= \sum_{n_k=0}^m e^{\omega(n-n_k)+3^{1/\beta-1} M \kappa_f n}. \end{aligned} \quad (27)$$

Theorem 2. Assume that G_1 and $G_3 - G_8$ are holds, then system (1) is β -Hyers–Ulam–Rassias stable over a discrete unbounded interval.

Proof. The solution of system (1) is

$$\begin{aligned} \Theta_n &= H^n \Theta_0 + H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \\ &\quad + \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}). \end{aligned} \quad (28)$$

Let Y_n be the solution of inequality (2), we have

$$\begin{aligned} &\left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right\| \\ &= \left\| \sum_{i=0}^{n-1} H^{n-1-i} h_i + \sum_{n_k=0}^n T(n-n_k) h_{n_k} \right\| \leq \sum_{i=0}^{n-1} \|H^{n-1-i}\| \|h\|_i + \sum_{n_k=0}^n \|T(n-n_k)\| \|h_{n_k}\| \\ &\leq M \left(\sum_{i=0}^{n-1} e^{\omega(n-i)} \varepsilon \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)} \varepsilon \Psi \right). \end{aligned} \quad (29)$$

Now, for each $n \in \{n_k, n_{k+1}\}$, we have

$$\begin{aligned} \|Y_n - \theta_n\|^\beta &= \left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \right. \\ &\quad \left. - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) \right\|^\beta \\ &= \left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \right. \\ &\quad \left. - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) \right. \\ &\quad \left. - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) + \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right. \\ &\quad \left. - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right\|^\beta \end{aligned}$$

$$\begin{aligned}
& \leq \left(\left\| \begin{aligned} & Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) \\ & - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \end{aligned} \right\| \right)^\beta + \left(\left\| H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) \right\| \right) \\
& - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \right)^\beta + \left(\left\| \begin{aligned} & \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \\ & - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) \end{aligned} \right\| \right)^\beta \\
& \leq \left(M \left(\sum_{i=0}^{n-1} e^{\omega(n-i)} \varepsilon \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)} \varepsilon \Psi \right) \right)^\beta + \left(M \sum_{i=0}^{n-1} e^{\omega(n-i)} \mathcal{L}_{f_i} \|Y_i - \Theta_i\| \right)^\beta \\
& + \left(M \sum_{n_k=0}^n e^{\omega(n-n_k)} \mathcal{L}_{I_{n_k}} \|Y_{n_k-1} - \Theta_{n_k-1}\| \right)^\beta, \tag{30}
\end{aligned}$$

if we set $\bar{Y}_n = e^{-\omega n} Y_n$ and $\bar{\Theta}_n = e^{-\omega n} \Theta_n$, then we have

$$\begin{aligned}
\|\bar{Y}_n - \bar{\Theta}_n\|^\beta & \leq \left(M \left(\sum_{i=0}^{n-1} e^{-\omega i} \varepsilon \Psi_i + \sum_{n_k=0}^m e^{-\omega n_k} \varepsilon \Psi \right) \right)^\beta + \left(M \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \|\bar{Y}_i - \bar{\Theta}_i\| \right)^\beta \\
& + \left(M \sum_{n_k=0}^n \mathcal{L}_{I_{n_k}} \|\bar{Y}_{n_k-1} - \bar{\Theta}_{n_k-1}\| \right)^\beta, \tag{31}
\end{aligned}$$

with the help of relation

$$(x + y + z)^\gamma \leq 3^{\gamma-1} (x^\gamma + y^\gamma + z^\gamma), \quad \text{where } x, y, z \geq 0 \tag{32}$$

$\gamma > 1.$

we get

$$\begin{aligned}
\|\bar{Y}_n - \bar{\Theta}_n\| & \leq 3^{(1/\beta)-1} M \left(\sum_{i=0}^{n-1} e^{-\omega i} \varepsilon \Psi_i + \sum_{n_k=0}^m e^{-\omega n_k} \varepsilon \Psi \right) + 3^{(1/\beta)-1} M \left(\sum_{i=0}^{n-1} \mathcal{L}_{f_i} \|\bar{Y}_i - \bar{\Theta}_i\| \right) \\
& + 3^{1/\beta-1} M \left(\sum_{n_k=0}^n \mathcal{L}_{I_{n_k}} \|\bar{Y}_{n_k-1} - \bar{\Theta}_{n_k-1}\| \right). \tag{33}
\end{aligned}$$

Using Lemma 1, we have

$$\|\bar{Y}_n - \bar{\Theta}_n\| \leq 3^{(1/\beta)-1} M \varepsilon \left(\sum_{i=0}^{n-1} e^{-\omega i} \Psi_i + \sum_{n_k=0}^m e^{-\omega n_k} \Psi \right) \left(1 + 3^{(1/\beta)-1} M \mathcal{L}_{I_{n_k}} \right)^k \exp \left(3^{(1/\beta)-1} M \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \right). \tag{34}$$

Resubmitting the values, we have

$$\begin{aligned}
\|Y_n - \Theta_n\| &\leq 3^{(1/\beta)-1} M \varepsilon \left(\sum_{i=0}^{n-1} e^{\omega(n-i)} \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)} \Psi \right) \left(1 + 3^{(1/\beta)-1} M \mathcal{L}_{I_{n_k}} \right)^k \exp \left(3^{(1/\beta)-1} M \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \right) \\
&\leq 3^{1/\beta-1} M \varepsilon \mathcal{L}_{\mathcal{F}} \left(\sum_{i=0}^{n-1} e^{\omega(n-i)} \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)} \Psi \right) \exp \left(3^{(1/\beta)-1} M (\kappa_f n + \varkappa_f) \right) \\
&= 3^{(1/\beta)-1} M \varepsilon \mathcal{L}_{\mathcal{F}} \left(\sum_{i=0}^{n-1} e^{\omega(n-i)+3^{(1/\beta)-1} M (\kappa_f n + \varkappa_f)} \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)+3^{(1/\beta)-1} M (\kappa_f n + \varkappa_f)} \Psi \right) \\
&= 3^{(1/\beta)-1} M \varepsilon \mathcal{L}_{\mathcal{F}} e^{3^{(1/\beta)-1} M \kappa_f} \left(\sum_{i=0}^{n-1} e^{\omega(n-i)+3^{(1/\beta)-1} M \kappa_f n} \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)+3^{(1/\beta)-1} M \kappa_f n} \Psi \right) \\
&\leq 3^{(1/\beta)-1} M \varepsilon \mathcal{L}_{\mathcal{F}} e^{3^{1/\beta-1} M \varkappa_f} (\eta_\varphi \varphi + M_1 \Psi) \\
&\leq 3^{(1/\beta)-1} M \varepsilon \mathcal{L}_{\mathcal{F}} e^{3^{(1/\beta)-1} M \varkappa_f} (M_1 + \eta_\varphi) (\varphi_n + \Psi) \\
\|Y_n - \Theta_n\|^\beta &\leq \mathcal{H}_{M, \mathcal{L}_{\mathcal{F}}, \eta_\varphi, M_1} \varepsilon^\beta (\varphi_n^\beta + \Psi^\beta),
\end{aligned} \tag{35}$$

where

$$\mathcal{H}_{M, \mathcal{L}_{\mathcal{F}}, \eta_\varphi, M_1} = 3^{1-\beta} \left(M \mathcal{L}_{\mathcal{F}} e^{3^{(1/\beta)-1} M \varkappa_f} (M_1 + \eta_\varphi) \right)^\beta > 0. \tag{36}$$

Thus, system (1) is β -Hyers–Ulam–Rassias stable. \square

Remark 3. Wang et al. [18] studied the β -Hyers–Ulam stability and β -Hyers–Ulam–Rassias stability for a system of impulsive differential equations as we know that difference equations relate to differential equations as discrete mathematics relate to continuous mathematics. The system of

impulsive difference equations used in this article is analogous to the system of impulsive differential equations used in [18]. Thus, the findings of this article are the discrete version of the work of Wang et al. [18].

6. Generalized β -Hyers–Ulam–Rassias Stability

In this section, we present the generalized β -Hyers–Ulam–Rassias stability, for which we need the following assumptions:

$$\begin{aligned}
G_9: 3^{(1/\beta)-1} M \varepsilon \prod_{n_k=0}^n \left(1 + 3^{(1/\beta)-1} M \mathcal{L}_{I_{n_k}} \right) \left(\sum_{i=0}^{n-1} e^{\omega(n-i)+3^{(1/\beta)-1} M \sum_{i=0}^{n-1} \mathcal{L}_{f_i}} \Psi_i \right) &\leq \eta_\varphi \varphi_n, \\
G_{10}: 3^{(1/\beta)-1} M \varepsilon \prod_{n_k=0}^n \left(1 + 3^{(1/\beta)-1} M \mathcal{L}_{I_{n_k}} \right) \left(\sum_{n_k=0}^m e^{\omega(n-n_k)+3^{(1/\beta)-1} M \sum_{i=0}^{n-1} \mathcal{L}_{f_i}} \Psi \right) &\leq \eta_\Psi \Psi_{k+1}.
\end{aligned} \tag{37}$$

Theorem 3. Assume that G_1 and $G_3 - G_{10}$ are satisfied, then system (1) is generalized β -Hyers–Ulam–Rassias stable.

Proof. The solution of system (1) is as follows:

$$\begin{aligned}
\Theta_n &= H^n \Theta_0 + H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \\
&\quad + \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}).
\end{aligned} \tag{38}$$

Let Y_n be the solution of inequality (2), we have

$$\begin{aligned}
& \left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^n H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^n H^{-i} f(i, Y_i, \zeta_i) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right\| \\
&= \left\| \sum_{i=0}^{n-1} H^{n-1-i} h_i + \sum_{n_k=0}^n T(n-n_k) h_{n_k} \right\| \leq \sum_{i=0}^{n-1} \|H^{n-1-i}\| \|h_i\| + \sum_{n_k=0}^n \|T(n-n_k)\| \|h_{n_k}\| \\
&\leq M \left(\sum_{i=0}^{n-1} e^{\omega(n-i)} \varepsilon \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)} \varepsilon \Psi \right).
\end{aligned} \tag{39}$$

Now, for each $n \in \{n_k, n_{k+1}\}$, we have

$$\begin{aligned}
\|Y_n - \theta_n\|^\beta &= \left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) \right\|^\beta \\
&= \left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) + H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) + \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right\|^\beta \\
&\leq \left(\left\| Y_n - H^n \Theta_0 - H^{n-1} \sum_{i=0}^{n-1} H^{-i} B \zeta_i - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) \right\| \right)^\beta \\
&\quad + \left(\left\| H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, Y_i, \zeta_i) - H^{n-1} \sum_{i=0}^{n-1} H^{-i} f(i, \Theta_i, \zeta_i) \right\| \right)^\beta \\
&\quad + \left(\left\| \sum_{n_k=0}^n T(n-n_k) I_k(n_k, Y_{n_k-1}, \zeta_{n_k-1}) - \sum_{n_k=0}^n T(n-n_k) I_k(n_k, \Theta_{n_k-1}, \zeta_{n_k-1}) \right\| \right)^\beta \\
&\leq \left(M \left(\sum_{i=0}^{n-1} e^{\omega(n-i)} \varepsilon \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)} \varepsilon \Psi \right) \right)^\beta + \left(M \sum_{i=0}^{n-1} e^{\omega(n-i)} \mathcal{L}_{f_i} \|Y_i - \Theta_i\| \right)^\beta \\
&\quad + \left(M \sum_{n_k=0}^n e^{\omega(n-n_k)} \mathcal{L}_{I_{n_k}} \|Y_{n_k-1} - \Theta_{n_k-1}\| \right)^\beta,
\end{aligned} \tag{40}$$

if we set $\bar{Y}_n = e^{-\omega n} Y_n$ and $\bar{\Theta}_n = e^{-\omega n} \Theta_n$, then we have

$$\begin{aligned} \|\bar{Y}_n - \bar{\Theta}_n\|^\beta &\leq \left(M \left(\sum_{i=0}^{n-1} e^{-\omega i} \varepsilon \Psi_i + \sum_{n_k=0}^m e^{-\omega n_k} \varepsilon \Psi \right) + \left(M \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \|\bar{Y}_i - \bar{\Theta}_i\| \right) \right)^\beta \\ &\quad + \left(M \sum_{n_k=0}^n \mathcal{L}_{I_{n_k}} \|\bar{Y}_{n_k-1} - \bar{\Theta}_{n_k-1}\| \right)^\beta, \end{aligned} \quad (41)$$

with the help of relation

$$(x + y + z)^\gamma \leq 3^{\gamma-1} (x^\gamma + y^\gamma + z^\gamma), \quad \text{where } x, y, z \geq 0 \quad (42)$$

$$\gamma > 1,$$

we get

$$\begin{aligned} \|\bar{Y}_n - \bar{\Theta}_n\| &\leq 3^{(1/\beta)-1} M \left(\sum_{i=0}^{n-1} e^{-\omega i} \varepsilon \Psi_i + \sum_{n_k=0}^m e^{-\omega n_k} \varepsilon \Psi \right) + 3^{(1/\beta)-1} M \left(\sum_{i=0}^{n-1} \mathcal{L}_{f_i} \|\bar{Y}_i - \bar{\Theta}_i\| \right) \\ &\quad + 3^{(1/\beta)-1} M \left(\sum_{n_k=0}^n \mathcal{L}_{I_{n_k}} \|\bar{Y}_{n_k-1} - \bar{\Theta}_{n_k-1}\| \right), \end{aligned} \quad (43)$$

using, we have

$$\|\bar{Y}_n - \bar{\Theta}_n\| \leq 3^{(1/\beta)-1} M \varepsilon \prod_{n_k=0}^n \left(\sum_{i=0}^{n-1} e^{-\omega i} \Psi_i + \sum_{n_k=0}^m e^{-\omega n_k} \Psi \right) \left(1 + 3^{(1/\beta)-1} M \mathcal{L}_{I_{n_k}} \right) \exp \left(3^{(1/\beta)-1} M \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \right), \quad (44)$$

resubmitting the values, we have

$$\begin{aligned} \|Y_n - \Theta_n\| &\leq 3^{(1/\beta)-1} M \varepsilon \prod_{n_k=0}^n \left(\sum_{i=0}^{n-1} e^{\omega(n-i)} \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)} \Psi \right) \left(1 + 3^{(1/\beta)-1} M \mathcal{L}_{I_{n_k}} \right) \exp \left(3^{(1/\beta)-1} M \sum_{i=0}^{n-1} \mathcal{L}_{f_i} \right) \\ &= 3^{(1/\beta)-1} M \varepsilon \prod_{n_k=0}^n \left(\sum_{i=0}^{n-1} e^{\omega(n-i)+3^{(1/\beta)-1} M \sum_{i=0}^{n-1} \mathcal{L}_{f_i}} \Psi_i + \sum_{n_k=0}^m e^{\omega(n-n_k)+3^{(1/\beta)-1} M \sum_{i=0}^{n-1} \mathcal{L}_{f_i}} \Psi \right) \left(1 + 3^{(1/\beta)-1} M \mathcal{L}_{I_{n_k}} \right), \end{aligned} \quad (45)$$

At last, we obtain

$$\begin{aligned} \|Y_n - \Theta_n\| &\leq 3^{1/\beta-1} M \varepsilon (\eta_\varphi \varphi_n + \eta_\Psi \Psi_{k+1}) \\ &\leq 3^{1/\beta-1} M \varepsilon (\eta_\varphi + \eta_\Psi) (\varphi_n + \Psi_{k+1}) \\ \|Y_n - \Theta_n\|^\beta &= 3^{1-\beta} M^\beta \varepsilon^\beta (\eta_\varphi + \eta_\Psi)^\beta (\varphi_n + \Psi_{k+1})^\beta \\ &= \mathcal{F}_{M, \eta_\varphi, \eta_\Psi, f} \varepsilon^\beta (\varphi_n^\beta + \Psi_{k+1}^\beta), \end{aligned} \quad (46)$$

where

$$\mathcal{F}_{M, \eta_\varphi, \eta_\Psi, f} = 3^{1-\beta} M^\beta (\eta_\varphi + \eta_\Psi)^\beta > 0. \quad (47)$$

Hence, system (1) is generalized β -Hyers-Ulam-Rassias stable. \square

7. Example

The impulsive difference system is as follows:

$$\begin{cases} \Theta_{n+1} = H\Theta_n + B\zeta_n + f(n, \Theta_n, \zeta_n), & n \in \{0, 1, 2, 3\}, \\ \Theta_0, & n = 0, \\ \Delta(\Theta_{n_k}) = I_k(n, \Theta_{n_{k-1}}, \zeta_{n_{k-1}}) = \frac{1}{3r^2}(\Theta_{n_{k-1}}), & k = 1, 2, 3, \dots, m. \end{cases} \quad (48)$$

Assumption G_1 and G_2 are holds if $N = 1$ and $M = 1$. Clearly, assumptions G_5 and G_6 are hold if $\kappa_f = 0$ and

$$\begin{aligned} \mathcal{L}_J &= \sup \prod_{r=1}^{\infty} \left(1 + 3^{(1/\beta)} M \mathcal{L}_{I_{n_k}} \right) \\ &= \sup \prod_{r=1}^{\infty} \left(1 + \frac{1}{r^2} \right) \leq e \sum_{r=1}^{\infty} (1/r^2) \leq e^{(\pi^2/6)}. \end{aligned} \quad (49)$$

Also, $\dot{u}_f = 0$. Now, set $\varphi_n = e^n$ and $\psi = 1$, then G_7 holds if $\eta_\phi = 1/2$. G_8 holds if $e^2/e^2 - 1$. Thus, system (3) is $1/2$ -Hyers-Ulam-Rassias stable with respect to $(\sqrt{e^n}, 1)$ on Z_+ and $\mathcal{H}_{M, \mathcal{L}_f, \eta_\phi, M_1} = \sqrt{3}e^{\pi^2/r^2} (1/2 + e^2/e^2 - 1)^{1/2}$.

8. Conclusion

Nowadays, studies on the qualitative behavior of impulsive difference equations have a significant contribution to the literature. In particular, the discussion regarding the β -Hyers-Ulam-Rassias stability of difference equations has been considered as one of the important topics of the literature, in which different types of conditions have been used in the form of inequalities, and most results have been obtained through discrete Gronwall inequality. In this paper, we have investigated the existence and uniqueness of the solution through the Banach contraction principle and β -Hyers-Ulam-Rassias stability of the impulsive difference system with the help of Gronwall inequality.

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.

Authors' Contributions

All authors contributed equally to this article and approved the final manuscript.

Acknowledgments

The Deanship of Scientific Research at King Khalid University, Abha, Saudi Arabia, has funded this work through Large Groups Project under grant number RGP.2/211/43.

References

- [1] N. U. Ahmed, "Existence of optimal controls for a general class of impulsive systems on Banach spaces," *SIAM Journal on Control and Optimization*, vol. 42, no. 2, pp. 669–685, 2003.
- [2] D. D. Bainov, V. Lakshmikantham, and P. S. Simeonov, "Theory of impulsive differential equations," *Series in Modern Applied Mathematics; World scientific*, vol. 6, 1989.
- [3] M. Benchohra, J. Henderson, and S. Ntouyas, *Impulsive differential equations and inclusions*, Hindawi Publishing Corporation, London UK, 2006.
- [4] E. Hernández, "On a new class of abstract impulsive differential equations," *Proceedings of the American Mathematical Society*, vol. 141, no. 5, pp. 1641–1649, 2012.
- [5] M. Pierri, "Existence of solutions for semi-linear abstract differential equations with not instantaneous impulses," *Applied Mathematics and Computation*, vol. 219, no. 12, pp. 6743–6749, 2013.
- [6] S. M. Ulam, *A Collection of Mathematical Problems*, Interscience Publishers, Hoboken New Jersey, USA, 1960.
- [7] D. H. Hyers, "On the stability of the linear functional equation," *Proceedings of the National Academy of Sciences*, vol. 27, no. 4, pp. 222–224, 1941.
- [8] T. Aoki, "On the stability of the linear transformation in Banach spaces," *Journal of the Mathematical Society of Japan*, vol. 2, no. 1-2, pp. 64–66, 1950.
- [9] T. M. Rassias, "On the stability of the linear mapping in Banach spaces," *Proceedings of the American Mathematical Society*, vol. 72, no. 2, pp. 297–300, 1978.
- [10] J. Wang, M. Feckan, and Y. Zhou, "Ulam's type stability of impulsive ordinary differential equations," *Journal of Mathematical Analysis and Applications*, vol. 395, no. 1, pp. 258–264, 2012.
- [11] J. Wang, Y. Zhou, and Z. Lin, "On a new class of impulsive fractional differential equations," *Applied Mathematics and Computation*, vol. 242, pp. 649–657, 2014.
- [12] A. Zada, O. Shah, and R. Shah, "Hyers-Ulam stability of non-autonomous systems in terms of boundedness of Cauchy problems," *Applied Mathematics and Computation*, vol. 271, pp. 512–518, 2015.
- [13] Y. Guo, M. Chen, X. B. Shu, and F. Xu, "The existence and Hyers-Ulam stability of solution for almost periodical fractional stochastic differential equation with fBm," *Stochastic Analysis and Applications*, vol. 39, no. 4, pp. 643–666, 2021.
- [14] Y. Guo, X. B. Shu, Y. Li, and F. Xu, "The existence and Hyers-Ulam stability of solution for an impulsive Riemann-Liouville fractional neutral functional stochastic differential equation with infinite delay of order $1 < \beta < 2$," *Boundary Value Problems*, vol. 1, p. 59, 2019.
- [15] G. Rahmat, A. Ullah, A. U. Rahman, M. Sarwar, T. Abdeljawad, and A. Mukheimer, "Hyers-Ulam stability of non-autonomous and nonsingular delay difference equations," *Advances in Difference Equations*, vol. 2021, p. 474, 2021.

- [16] W. Hu and Q. Zhu, "Stability criteria for impulsive stochastic functional differential systems with distributed-delay dependent impulsive effects," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 3, pp. 1–6, 2019.
- [17] W. Hu, Q. Zhu, and H. R. Karimi, "Some improved Razumikhin stability criteria for impulsive stochastic delay differential systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 12, pp. 5207–5213, 2019.
- [18] X. Wang, M. Arif, and A. Zada, " β -Hyers-Ulam-Rassias stability of semilinear nonautonomous impulsive system," *Symmetry*, vol. 11, no. 2, p. 231, 2019.

Research Article

Key Information Extraction Algorithm of Different Types of Digital Archives for Cultural Operation and Management

Xiulun Ma 

University of Jinan, Jinan 250022, China

Correspondence should be addressed to Xiulun Ma; shc_maxl@ujn.edu.cn

Received 4 July 2022; Accepted 11 August 2022; Published 29 August 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiulun Ma. 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 effect of key information extraction from digital archives, a key information extraction algorithm for different types of digital archives is designed. Preprocess digital archive information, taking part of speech and marks as key information. Self-organizing feature mapping network is used to extract the key information features of digital archives, and the semantic similarity calculation results are obtained by combining the feature extraction results. Combine with mutual information collection, take that word with the highest mutual information value as the collection cent, traverse all keywords, and take the central word as the key information of digital archives to complete the extraction of key information. Experiments show that the recall rate of the algorithm ranges from 96% to 99%, the extraction accuracy of key information of digital archives is between 96 and 98%, and the average extraction time of key information of digital archives is 0.63 s. The practical application effect is good.

1. Introduction

Generally speaking, there are no management problems in the spontaneous stage of cultural production. Cultural production and business activities are the product of commodity economy [1, 2]. When material and cultural production develop to a certain extent, social division of labor is further clarified, and professionals and professional groups engaged in cultural production appear, both the ruling class and the ruled class try to use cultural production to serve the interests of their own class, which is the conscious stage of cultural production [3]. Only at the conscious stage of cultural production can cultural operation and management be put on the agenda. In a modern capitalist society, everything into a commodity culture products without exception has become a part of capitalists, for-profit special goods. The vast majority of cultural activities are restricted by the value of commodity production rules [4]. The tendency of commercialization of cultural production and cultural activities has become a common social phenomenon in the field of capitalist culture. The law of market economy dominates the management of cultural operations and management activities, and the quality of management is the key to the success or failure of specific cultural

products in the free competition of the cultural market. Especially with the rapid development of social economy, different types of digital archives are gradually increasing, the main products of these are social culture, in order to better manage for these different types of digital files, need to study a new key to different types of digital archives information extraction algorithm, to enhance the management level of the digital scheme. Therefore, it is of great significance to study a key information extraction algorithm of different types of digital archives.

In the digital archives, information extraction is an important research topic, reference [5] proposed a digital book records mass data fast extraction algorithm. Based on the range characteristics of large-scale data attributes, the distribution samples of digital book archive data are divided into multiple subintervals to achieve data classification. By constructing a neuron model, the error terms of output are determined according to the data output of the hidden layer and output layer, and the weight of each layer of the BP neural network is adjusted. This method builds a fast extraction model based on BP neural network and realizes the fast extraction of massive archive data. Reference [6] proposed a key information extraction algorithm based on TextRank and cluster filtering. First, the key information is

extracted and vectorized for Word2Vec. Then, TextRank is improved by constructing a graph model integrating word eigenvalues and edge weights, and the stable graphs obtained by iterative convergence are merged and clustered to form clusters. Then, a cluster quality evaluation formula was designed for cluster filtering, and TextRank was applied to form the final clustering. Finally, annotate the information type of the cluster. For testing the text, by comparing the key information vector distance cluster heart vector and the words information types, combines information type and key information to get the key of the text information. Reference [7] proposed a hidden Markov model based on an improved extraction algorithm of key information extraction. The web document is converted into DOM tree and preprocessed, and the information item to be extracted is mapped to state and the observation item to be extracted is mapped to vocabulary. The improved hidden Markov model is used to extract key information of the text. Reference [8] proposed a key information extraction algorithm based on word vector and location information. Vector representation model by word learning vector of each word in the target document said, will the reflect of the latent semantic relations between the word and the word vector combined with location feature fusion to the PageRank score model, choose a few top words or phrases as the key target document information, in order to complete the digital archives of key information extraction. Reference [9] proposed a key information extraction algorithm of unstructured text in the knowledge database. Six yuan group was used to optimize the hidden Markov model, probability model, and smooth processing of incomplete training samples. Initialization and termination operations were carried out for the sequences of observation values released at different times to obtain the optimal state sequence. After decoding the observation sequence, the positive sequence and reverse sequence were obtained by comparing them to filter out the states without decoding ambiguity and complete ambiguity elimination. According to the maximum probability state sequence, the text key information to be extracted is defined and the key information is extracted.

However, the above-mentioned key information extraction algorithm is suitable for different types of digital files, and the effect is not ideal because the boundary of key information extraction is uncertain. Therefore, this paper designs a new key information extraction algorithm for different types of digital archives. Firstly, the algorithm divides the main categories of key information, takes parts of speech and marks as features, and introduces the self-organizing feature mapping neural network to traverse the center of word set, thus realizing the extraction of key signals quickly and accurately. The effectiveness of the algorithm is verified by experiments.

2. Materials and Methods

2.1. Digital Archive Processing

2.1.1. The Text Participle. In the process of cultural management, there are many types of digital archives. Before

extracting the key information of different types of digital archives, it is necessary to preprocess the key information of digital archives. The preprocessing process includes word segmentation and marking. Word segmentation refers to classifying the words in the text and setting the marks according to the categories, which lays the foundation for the key information extraction of digital archives in the future [10].

The difference between the word segmentation process of the reverse maximum matching algorithm and the forward maximum matching algorithm is that the scanning of the reverse maximum matching algorithm starts from the end of the string. Each unsuccessful match removes the preceding word until the match is successful. Then the basic idea of the bidirectional maximum matching algorithm is: When segmenting different types of digital archives information, firstly, a forward word-for-word maximum matching algorithm is applied to the character string to be processed, then a reverse word-for-word maximum matching algorithm is applied, and the output result is used to complete the word segmentation processing. Assuming bidirectional maximum matching word segmentation for $S = (C_1, C_2, \dots, C_i)$, the algorithm process can be described as follows:

- (1) First take out the first word C_1 in S , and search in the dictionary to see if there are any words with C_1 as the prefix. If there are, save them as word marks [11].
- (2) Take a word C_2 from S and match it with the dictionary to determine whether there is a word with C_2 as the prefix.
- (3) If it does not exist, split C_1 from string S , ending with a word split.
- (4) If there is, to determine whether C_1C_2 into words, calculate the number n headed by C_1C_2 words.
- (5) If $n = 0$, the participle ends once [12].
- (6) If n is not 0, then take a word C_i from S and match it with the dictionary to determine whether there is a word prefixed with C_1, C_2, \dots, C_i .
- (7) If yes, go to Step 6.
- (8) If it does not exist, split C_1, C_2, \dots, C_{i-1} from string S , ending with a word split.
- (9) Continue word segmentation from string C_i of S , repeat the above steps until the end of string S forward segmentation.
- (10) Take out the last word C_n in S and match it in the dictionary to find whether there is a word with suffix C_1 . If so, save it as a word mark [13].
- (11) Then take out a word C_{n-1} from S and match it with the dictionary to judge whether there is a word with suffix C_1C_2 .
- (12) If it does not exist, it splits C_n from string S , ending with a word split.
- (13) If there is, then judge whether $C_{n-1}C_n$ is a word and count the number of words starting with $C_{n-1}C_n$, expressed by n .

- (14) If $n = 0$, then the participle ends.
- (15) If n is not 0, take out a word C_i from S and match it with the dictionary to determine whether there is a word with $C_i, \dots, C_{n-1}C_n$ as the suffix.
- (16) If yes, go to Step (15).
- (17) If it does not exist, $C_i, \dots, C_{n-1}C_n$ will be cut out from string S and a word segmentation will end.
- (18) Continue word segmentation from word C_i of string S , and repeat the above steps until the end of reverse segmentation of string S , so as to remove the stop word. The specific implementation process is shown in Figure 1.

2.1.2. The Part of Speech Tagging. Part of speech is a grammatical attribute of vocabulary, which generally indicates the type of a word in the corpus. Part-of-speech tagging refers to the process and method of tagging the part of speech of each word. Some words contain multiple parts of speech, with different parts of speech and completely different ways of expression [14, 15]. However, in general, when a word contains one or more parts of speech, the frequency of its commonly used parts of speech is far greater than that of other parts of speech, so the accuracy of POS tagging can be ensured on the whole, and the POS tagging method can be applied to most application scenarios [16]. Conditional Random Field Algorithm (CRF) was proposed by Lafferty et al. in 2001. It is an undirected graph model combining the characteristics of the maximum entropy model and hidden Markov model. In recent years, good results have been achieved in sequence tagging tasks such as word segmentation, part-of-speech tagging, and named entity recognition [17]. One of the simplest conditional random fields is the chain structure, in this special conditional random field, the chain structure is composed of several character marks. In CRF models with only one order chain, the fully connected subgraph covers the set of the current marker and one marker before it, as well as the maximum connected graph of any subset of the observation sequence. The chained conditional random field is shown in Figure 2, and the set of vertices can be regarded as the maximum connected subgraph.

In the sequence labeling task, random variable $X = \{x_1, x_2, \dots, x_n\}$ represents the observable sequence, random variable $Y = \{y_1, y_2, \dots, y_n\}$ represents the corresponding marker sequence of the observed sequence [18], and the chained conditional probability distribution of the random variable Y is:

$$p(y|x) = \frac{1}{Z(x)} \exp \left(\sum_{i,k} \lambda_k f_k(y_{i-1}, y_i, x) + \sum_{i,k} \lambda'_k f'_k(y_i, x) \right). \quad (1)$$

In the above formula, $f_k(y_{i-1}, y_i, x)$ is the state feature function for edge and capture mark transfer features. $\sum_{i,k} \lambda'_k f'_k(y_i, x)$ is the non-negative factor for each node. $f'_k(y_i, x)$ is the state feature function that captures the current marked feature for the edge. λ_k and λ'_k are learning model

parameters [19], said the weight of characteristic function. $Z(x)$ is a normalizing factor dependent only on the observation sequence. The specific calculation formula is as follows:

$$Z(x) = \exp \left(\sum_i \sum_k \lambda_k f_k(y_{i-1}, y_i, x) \right). \quad (2)$$

Conditional random field reasoning refers to finding a marker sequence $Y = \{y_1, y_2, \dots, y_n\}$ corresponding to the most probable one given an observation sequence $X = \{x_1, x_2, \dots, x_n\}$. In the distribution function of conditional random fields, the normalized factor is completely independent of the marker sequence [20]. Therefore, given the model parameters, the most likely marker sequence can be expressed as:

$$\begin{aligned} Y^* &= \operatorname{argmax}_y p(y|x) \\ &= \operatorname{argmax}_y \exp \left(\sum_{i,k} \lambda_k f_k(y_{i-1}, y_i, x) + \sum_{i,k} \lambda'_k f'_k(y_i, x) \right). \end{aligned} \quad (3)$$

When the current sequence position is i and the current label is y , the algorithm can be used to obtain the unnormalized probability value of the optimal label sequence to the current position. Its recursive form is:

$$\theta(i, y) = \max_{y'} \left\{ \theta \left(i-1, y' * l \sum_k \lambda_k f_k(x, y, y', i) \right) \right\}. \quad (4)$$

2.2. Key Information Feature Extraction of Digital Archives. Self-organizing feature mapping neural network was proposed by a professor of neural network expert self-organizing feature mapping network of University of Helsinki, Finland in 1981 [21]. This network simulates the function of self-organizing feature mapping of the brain nervous system. It is a kind of competitive learning network, which can carry out self-organizing learning without supervision in learning [22]. This paper uses this method to extract the key information features of different types of digital archives. This can improve the accuracy and efficiency of extracting key information from archives.

The structure of self-organizing feature mapping neural network is shown in Figure 3.

We set the number of neurons in the input layer to be n , and the number of neurons in the competition layer to be $M = m^2$. The input layer and the competition layer form a two-dimensional planar array. The two layers are connected, and sometimes neurons in the competing layer are also connected by edge inhibition [23]. There are two kinds of connection weights in the network, one is the connection weights of neurons responding to external inputs, and the other is the connection weights between neurons, whose size controls the size of interactions between neurons [24, 25].

The connections of neurons at the competitive layer of each input neuron in the self-organizing feature mapping

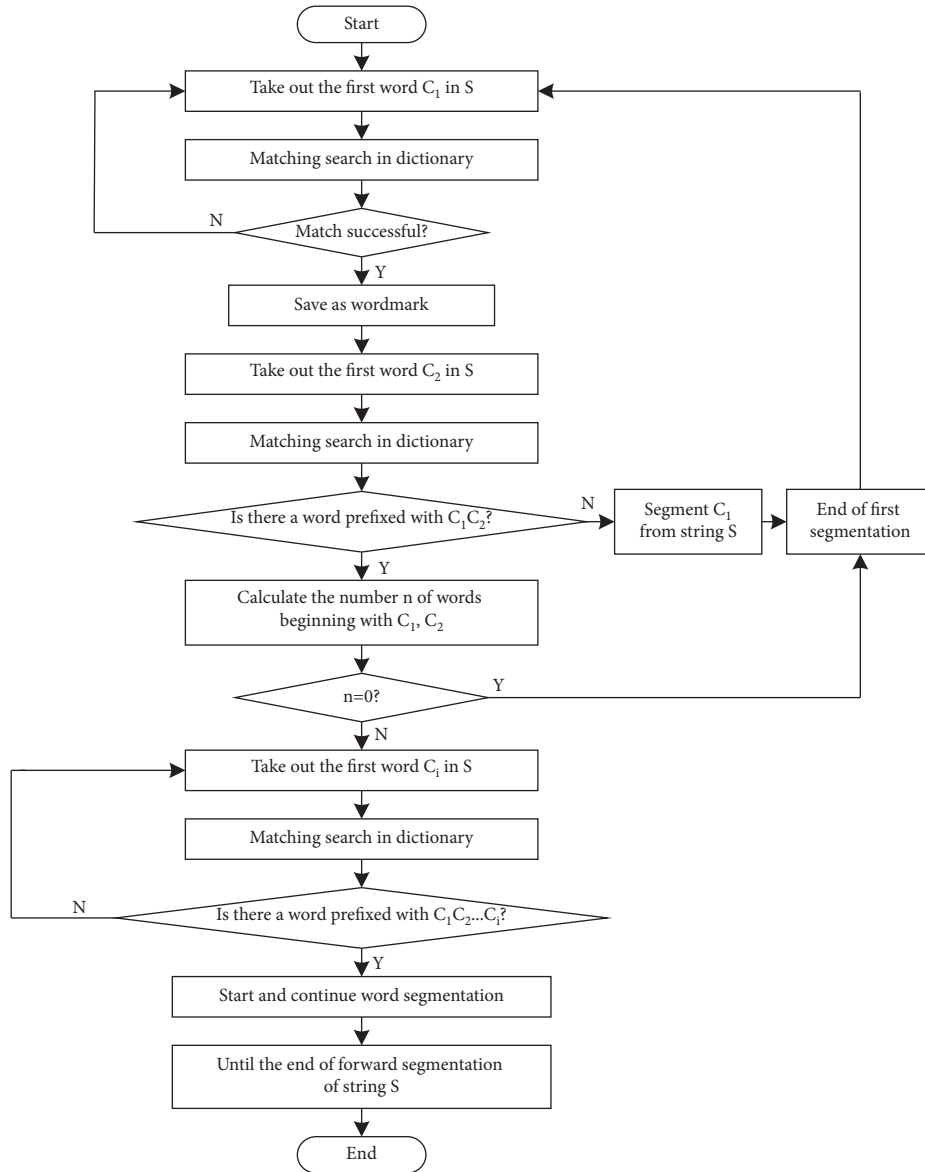


FIGURE 1: Text word segmentation and the process of removing stop words.

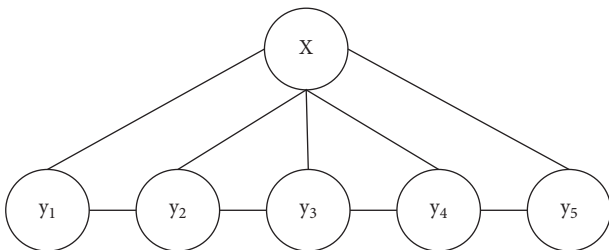


FIGURE 2: Chained conditional random fields.

network structure shown in Figure 3 are extracted, as shown in Figure 4.

Set the input mode of the network as $P_k = (p_1^k, p_2^k, \dots, p_n^k), k = 1, 2, \dots, q$ and the neuron vector of the competition layer as $A_j = (a_{j1}, a_{j2}, \dots, a_{jm}), j = 1, 2, \dots, m$. Where P_k is a continuous value and A_j is a

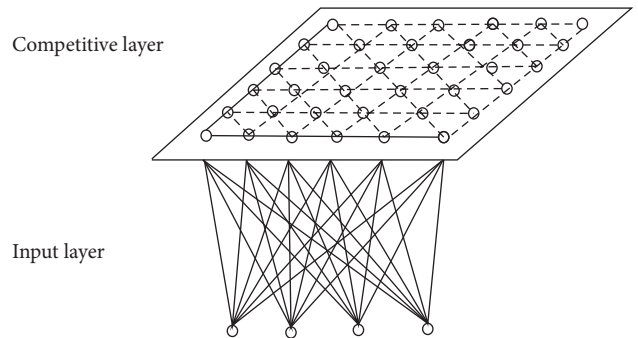


FIGURE 3: Structure of self-organizing feature mapping neural network.

numerical quantity. The connection vector between neuron j of the competition layer and neuron of the input layer is $W_j = (w_{j1}, w_{j2}, \dots, w_{jm}), j = 1, 2, \dots, M$.

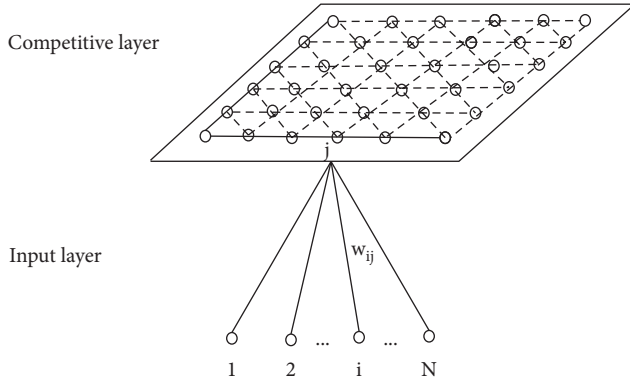


FIGURE 4: Self-organizing feature mapping network structure after extracting competitive layer neurons.

The self-organizing learning process of the self-organizing feature mapping network can also be described as: for each input of the network, only part of the weight is adjusted to make the weight vector closer to or more deviated from the input vector. This adjustment process is competitive learning. With continuous learning, ownership vectors are separated from each other in vector space, forming a class of patterns representing input space, respectively, which is the clustering function of automatic feature recognition in a self-organizing feature mapping network. The learning and working rules of the network are as follows:

(1) Initialization

Assign the network connection weight $\{w_{ij}\}$ to the random value $i = 1, 2, \dots, N$, $j = 1, 2, \dots, M$ in the interval $[0, 1]$. The initial value of learning rate $\eta(t)$, $0 < \eta(t) < 1$ was determined. Determine the initial value $N_g(0)$ of neighborhood $N_g(t)$. Neighborhood $N_g(t)$ is essentially a region centered on the winning neuron g and contains several neurons. This area is generally uniformly symmetrical, most typically a square or circular area. The value of $N_g(t)$ represents the number of neurons in the neighborhood during the t -th learning. Determine the total number of studies T .

(2) One of the q learning modes P_k, P_k is provided to the input layer of the network and normalized. The specific calculation formula is as follows:

$$\bar{P}_k = \frac{P_k}{P_k} = \frac{(p_1^k, p_2^k, \dots, p_n^k)}{\left[(p_1^k)^2 + (p_2^k)^2 + \dots + (p_n^k)^2 \right]^{1/2}} \quad (5)$$

(3) Normalize the connection weight vector $W_j = (w_{j1}, w_{j2}, \dots, w_{jN})$ and calculate the Euclidean distance between \bar{W}_j and \bar{P}_k . The calculation formula of \bar{W}_j is as follows:

$$\bar{W}_j = \frac{W_j}{W_j} = \frac{(w_1^k, w_2^k, \dots, w_n^k)}{\left[(w_1^k)^2 + (w_2^k)^2 + \dots + (w_n^k)^2 \right]^{1/2}} \quad (6)$$

The Euclidean distance between \bar{W}_j and \bar{P}_k can be calculated by the following formula:

$$d_j = \left[\sum_{i=1}^N (\bar{P}_i^k - \bar{W}_i^k)^2 \right], \quad j = 1, 2, \dots, M. \quad (7)$$

(4) Find the minimum distance d_g and determine the winning neuron g .

$$d_g = \min[d_j]. \quad (8)$$

(5) Adjust the connection weights, and modify the connection weights between all neurons in neighborhood $N_g(t)$ of the competition layer and neurons of the input layer. The specific formula is as follows:

$$\bar{w}_{ji}(t+1) = \bar{w}_{ji}(t) + \eta(t) \left[\bar{P}_i^k - \bar{w}_{ji}(t) \right]. \quad (9)$$

In the above formula, $\eta(t)$ is the learning rate at moment t .

(6) Select another learning mode to provide to the input layer of the network and return to step (3) until all q learning modes are provided to the network.

(7) Updated learning rate $\eta(t)$ and neighborhood $N_g(t)$.

$$\eta(t) = \eta(0) \left(1 - \frac{1}{T} \right). \quad (10)$$

In the above formula, $\eta(0)$ is the initial learning rate, t is the number of learning, and T is the total number of learning.

Assume that the coordinate value of a certain neuron g in the competition layer in the two-dimensional array is (x_g, y_g) , then the range of neighborhood is point $(x_g + N_g(t), y_g + N_g(t))$ and point $(x_g - N_g(t), y_g - N_g(t))$ as the square in the upper right corner and the lower left corner, and the modified formula is as follows:

$$N_g(t) = \text{INT} \left[N_g(0) \left(1 - \frac{1}{T} \right) \right]. \quad (11)$$

In the above formula, $\text{INT}(\cdot)$ is the integral function.

(8) Let $t = t + 1$, return to step (2), until $t = T$.

2.3. Key Information Extraction Algorithm of Digital Archives.

Key in the process of information extraction, in the digital archives to effectively extract the digital archives of key information, cannot individually understand the individual words of digital archives, and words or similar to each other in the digital archives correlation words combined into a block, a comprehensive understanding of the whole text content and the exact meaning of each word. Therefore, the semantic similarity between words is used as the clustering distance. All the semantemes of a word will form a hierarchical structure similar to a tree according to their upper and lower positional relations, which is traversed through the tree. Finally, the distance between words can be used to judge the similarity of word meaning. The formula for calculating word distance is as follows:

$$\text{Sim}(p_1, p_2) = \frac{\alpha}{\alpha + \text{dist}(p_1, p_2)}. \quad (12)$$

In the above formula, p_1 and p_2 represent two semesters, which are variable parameters. $\text{dist}(p_1, p_2)$ represents the length of the path between two sememes of a word. The semantic origin of describing concepts is divided into four parts: The first basic semantic origin, the symbolic semantic origin, the relational semantic origin, and other independent semantic origin. The overall similarity between concepts is calculated by the following formula:

$$\text{Sim}(s_1, s_2) = \sum_{i=1}^n y_i \prod_{j=1}^i \text{Sim}_j(p_1, p_2). \quad (13)$$

In the above formula, s_1 and s_2 represent two concepts, and y_i represents the result of feature extraction. If there are two words w_1 and w_2 in the set, among which word w_1 has n concept descriptions and word w_2 has m concept descriptions, the maximum similarity between concepts w_1 and w_2 can be used as the semantic similarity of the two words, and the calculation formula is as follows:

$$\text{Sim}(w_1, w_2) = \max_{i=1, \dots, n, j=1, \dots, m} \text{Sim}(s_{1i}, s_{2j}). \quad (14)$$

The process of key information extraction algorithm of digital archives is as follows:

Preprocessing: Word segmentation for digital archival text, stop word overconsideration.

Step 1: Calculate all candidate words and semantic similarities between w_i and w_j in digital archival text $\text{Sim}(w_i, w_j)$.

(1) TF-IDF value is calculated, and word $W = \{W_1, W_2, \dots, W_N\}$ with word frequency greater than the threshold t is selected as the candidate key information. The calculation formula of TF-IDF value is as follows:

$$\frac{TF}{IDF_i} = \frac{tf_i \times \log(N/n_i)}{\sqrt{\sum_j (tf_j \times \log(N/n_j))^2}} \quad (15)$$

In the above formula, tf_i is the number of occurrences of the word in the current digitized archival text, N is the total number of digitized archival text, and n_i is the number of digitized archives containing the word w_i in the database.

- (2) During initialization, each word $\{W_i\}$ in the candidate word has a cluster Z_i , a total of n clusters, and all of them are set with unaccessed markers.
- (3) Among all non-visited word clusters, select the cluster pair (C_l, C_k) with the largest similarity, that is, the closest distance, by calculating the maximum value of $\text{Sim}(w_i, w_j)$. If $\text{Sim}(C_l, C_k)$ is less than the given threshold, turn to (6); otherwise, merged clusters C_l and C_k are new clusters $C_0 = C_l \cup C_k$. Set to current cluster C , C to no access flag, C_l and C_k to access flag.
- (4) Calculate the semantic similarity among all unaccessed word clusters, and transfer to (4).

(5) After clustering, the first k words with better quality are selected from each cluster Z_i as the final key information, so as to obtain the candidate word set $W = \{C_1, C_2, \dots, C_m\}$.

Step 2: Treat each word in the text as a set C_j , a total of N sets (N is the number of words in the text).

Step 3: Select the two sets C_i and C_j with the greatest similarity from the N sets, and combine the two sets into a new set C .

Step 4: Select the center point of the current set: calculate the mutual information sum of the words in the current set and other words outside the set, and select the word with the largest mutual information value as the center point of the current set. If the calculated mutual information value between words is large, it indicates that they are also relatively large, on the contrary, it indicates that they are relatively small. The mutual information between w_i and w_j , that is, the public information between w_i and w_j , is calculated as follows:

$$I(w_i, w_j) = \log \frac{p(w_i|w_j)}{p(w_i)} = \log \frac{p(w_i, w_j)}{p(w_i)p(w_j)}. \quad (16)$$

In the above formula, $p(w_i, w_j)$ is the common frequency of w_i and w_j , $p(w_i)$ is the separate frequency of w_i , and $p(w_j)$ is the separate frequency of w_j . According to the above formula, when $I(w_i, w_j) > 0$, the greater the value, the more public information between w_i and w_j and the stronger the correlation; when $I(w_i, w_j) = 0$, there is less public information between w_i and w_j and the correlation is weak; when $I(w_i, w_j) < 0$, there is no correlation between w_i and w_j .

Step 5: Among other words outside the set, select the word with the highest similarity with the center point of the set. If the similarity value is greater than the threshold, add it to the current set C ; calculate the mutual information between the central point of the current set and the words outside the set, and add the word with the largest mutual information value to the current set C .

Step 6: Turn to step 4 to update the current collection center point until all words are accessed. If the mutual information value between the central point of the set and other words outside the set is less than 0, perform step 3 for the remaining unreachable words until all the words are accessed and divided.

Step 7: In the final cluster set, select its first K central words as the key information of the text. The key information extraction algorithm flow of different types of digital archives is shown in Figure 5.

3. Results and Discussion

3.1. Experimental Scheme. In order to verify the effectiveness of the algorithm designed in this paper to extract archive information, we conducted simulation experiments. This experiment is a simulation experiment, so it is necessary to

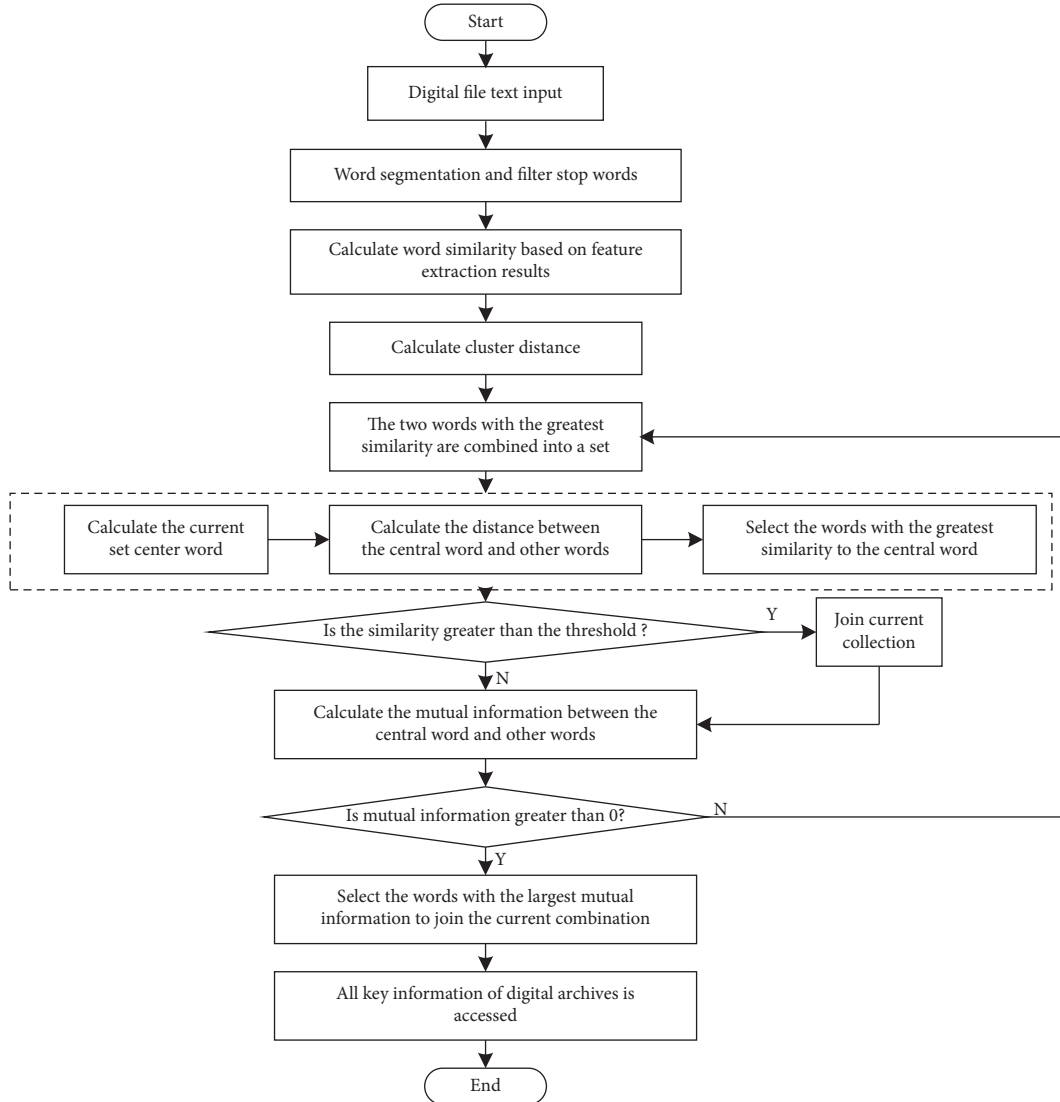


FIGURE 5: . Key information extraction algorithm flow of different types of digital archives.

design the experimental parameters, consider various factors, compare various types of simulation software and computers, and complete the design of environmental parameters of the simulation experiment, as shown in Table 1.

During the experiment, 500 GB digital archives were randomly selected from schools, enterprises, and relevant administrative units as data sets, and 450 GB of them were randomly selected as training sets to train this method. The remaining 50 GB were used as test sets to test the key information extraction performance of different types of digital archives. In order to ensure the objectivity of the experiment, the title and core prompt will be filtered out in the process of extracting key information. Recall rate and accuracy rate are often used as indicators of the key information extraction effect of different types of digital archives. Recall rate R and accuracy rate P adopted in this experiment are defined as follows:

$$R = \frac{j}{l} \times 100\%. \quad (17)$$

In the above formula, l represents the number of extracted key information, and j represents the actual number of key information.

$$P = \frac{L}{l} \times 100\%. \quad (18)$$

In the above formula, L represents the amount of key information accurately extracted.

The time-consuming calculation formula for extracting key information of different types of digital archives is as follows:

$$T = \sum_{i=1}^n t_i. \quad (19)$$

In the above formula, t_i represents the time taken for the i -th key information extraction step of digital archives.

3.2. Analysis and Discussion of Experimental Results. The recall rates of key information extraction of different types of

TABLE 1: Experimental environment parameters.

Experimental environment parameters	Configuration	Parameter
Hardware environment	CPU	Intel (R)Core (TM)i5-9400
	Frequency	2.90 GHz
	RAM	16.0 GB
Software environment	Operating system	Windows 10
	Analog software language	APDL
	Simulation software	Matlab 7.2

digital archives of reference [5] algorithm, reference [6] algorithm, reference [7] algorithm, and algorithm of this paper are compared. The results are shown in Figure 6.

By analyzing the data in Figure 6, we can see that the recall rate of the algorithm in reference [5] changes in the range of 58%–85%, the recall rate of the algorithm in reference [6] changes in the range of 49%–79%, and the recall rate of the algorithm in reference [7] changes in the range of 50%–87%. Compared with the experimental comparison algorithm, the recall rate of the algorithm of this paper changes in the range of 96%–99%, which is always higher than the experimental comparison algorithm, it shows that the key information of digital archives can be extracted comprehensively by using this algorithm, and the integrity is higher.

The key information extraction accuracy of different types of digital archives of reference [5] algorithm, reference [6] algorithm, reference [7] algorithm, and algorithm of this paper are compared. The results are shown in Figure 7.

By analyzing the data in Figure 7, we can see that the extraction accuracy of key information of digital archives of reference [5] algorithm is 49%–85%, the extraction accuracy of key information of digital archives of reference [6] algorithm is 54%–80%, and the extraction accuracy of key information of digital archives of reference [7] algorithm is 56%–80%. Compared with these algorithms, the extraction accuracy of key information of digital archives of the algorithm of this paper is 96%–98%. On the whole, the key information extraction accuracy of this algorithm is relatively stable, and there is no fluctuation of too high or too low, which indicates that the reliability of this algorithm in extracting key information is high. The accuracy of information extraction is higher, which can achieve the ultimate goal of accurately extracting the key information of different digital archives.

The extraction time of key information of different types of digital archives of reference [5] algorithm, reference [6] algorithm, reference [7] algorithm, and algorithm of this paper are compared. The comparison results are shown in Table 2.

By analyzing the results in Table 2, it can be seen that the average time-consuming of digital archives key information extraction of reference [5] algorithm is 1.41 s, the average time-consuming of digital archives key information extraction of reference [6] algorithm is 1.39 s, and the average time-consuming of digital archives key information extraction of reference [7] algorithm is 1.49 s, which is the

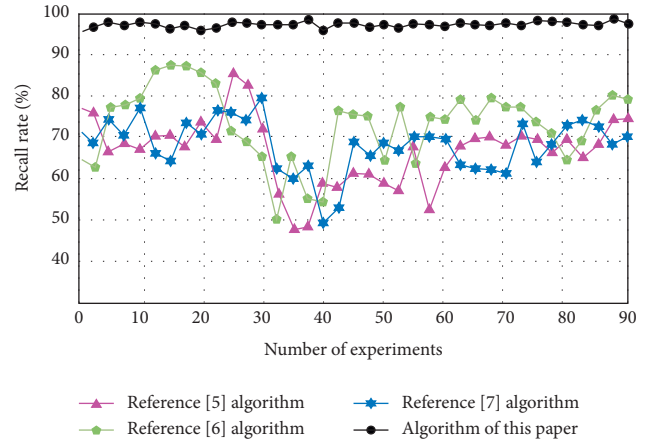


FIGURE 6: Comparison of recall rate.

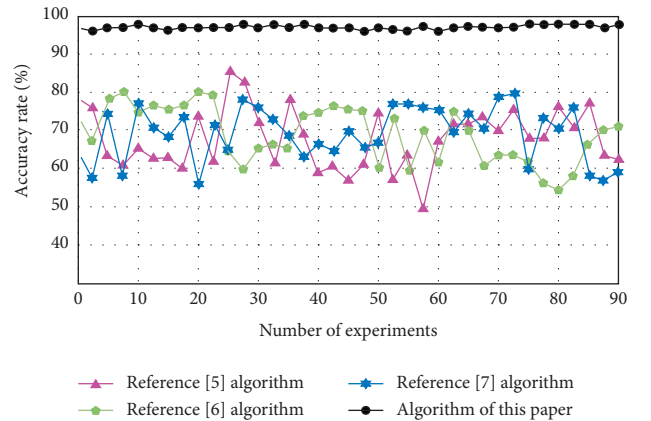


FIGURE 7: Comparison of accuracy.

highest among the four algorithms. Compared with these algorithms, the average extraction time of key information of digital archives in this algorithm is 0.63 s, which has a shorter extraction time and higher efficiency, and can realize the rapid extraction of key information of digital archives.

To sum up, the recall rate of this algorithm changes in the range of 96%–99%, the accuracy of key information extraction of digital archives is 96%–98%, and the average time-consuming of key information extraction of digital archives is 0.63 s. It can achieve the goal of rapid and accurate extraction of key information of digital archives, solve a variety of problems existing in traditional methods, and can be widely used in many fields.

TABLE 2: Extraction time of key information of different types of digital archives.

Number of experiments	Time (s)			
	Reference [5] algorithm	Reference [6] algorithm	Reference [7] algorithm	Algorithm of this paper
10	1.25	1.33	1.47	0.47
20	1.44	1.25	1.58	0.56
30	1.23	0.96	1.56	0.58
40	1.38	1.47	1.47	0.62
50	1.45	1.58	1.52	0.85
60	1.63	1.65	1.35	0.67
70	1.45	1.58	1.47	0.81
80	1.56	1.40	1.62	0.57
90	1.29	1.31	1.41	0.51
Average value	1.41	1.39	1.49	0.63

4. Conclusions

With the continuous optimization of cultural operation and management strategies, the level of cultural operation and management has been gradually improved, and digital archives management is an important part of cultural operation and management. Therefore, extracting the key information of different types of digital archives is of great significance to the level of cultural operation and management. Therefore, this paper designs a key information extraction algorithm of different types of digital archives for cultural operation and management. The experimental results show that the recall rate of the algorithm is between 96% and 99%, the accuracy of key information extraction of digital archives is 96%–98%, and the average time-consuming of key information extraction of digital archives is 0.63 s. It can achieve the goal of rapid and accurate extraction of key information of digital archives and can be widely used in cultural operation and management, in order to improve the quality of cultural operation and management to the greatest extent, promote the further development of the cultural industry. However, the convergence of this algorithm is not tested in the process of operation. In order to avoid falling into the local optimum, it is necessary to increase the optimization of the algorithm in future research work to avoid too many iterations or high errors.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] E. Ruikyte, "Reflections on "Then & Now": arts and cultural management and the shortcomings of student-led research projects," *Exchanges The Interdisciplinary Research Journal*, vol. 8, no. 4, pp. 87–98, 2021.
- [2] N. P. Gomes and W. A. Cantú, "Sociocultural trend reports as an intelligence tool of strategic cultural management," *Marketing and Smart Technologies*, vol. 191, no. 1, pp. 63–74, 2022.
- [3] Y. T. Lee and N. Y. A. Gyamfi, "The Sage handbook of contemporary cross-cultural management," *Journal of International Business Studies*, vol. 53, no. 3, pp. 568–571, 2022.
- [4] T. Jackson, "Book review: cases in critical cross-cultural management: an intersectional approach to culture," *International Journal of Cross Cultural Management*, vol. 21, no. 2, pp. 398–399, 2021.
- [5] L. Zhang, "Simulation Research on rapid extraction of massive data from digital books and archives," *Computer Simulation*, vol. 36, no. 3, pp. 397–400, 2019.
- [6] Z. B. Chen, Y. M. Li, and F. Xu, "Research on forestry text key information extraction based on textrank and cluster filtering," *Journal of Agricultural Machinery*, vol. 51, no. 5, pp. 207–214, 172.
- [7] Z. Q. Liu, Y. C. Du, and S. C. Shi, "Key information extraction of web news based on Improved Hidden Markov model," *Modern Library and Information Technology*, vol. 3, no. 3, pp. 120–128, 2019.
- [8] W. Fan, H. Liu, and Y. X. Zhang, "Key information extraction algorithm integrating word vector and position information," *Computer Engineering and Application*, vol. 56, no. 5, pp. 179–185, 2020.
- [9] W. J. Guo and X. A. Bao, "Key information extraction model of unstructured text in knowledge database," *Computer Simulation*, vol. 38, no. 9, pp. 357–360, 2021.
- [10] B. Rajyagor and R. Rakholia, "Tri-level handwritten text segmentation techniques for Gujarati language," *Indian Journal of Science and Technology*, vol. 14, no. 7, pp. 618–627, 2021.
- [11] B. Geng, "Text segmentation for patent claim simplification via bidirectional long-short term memory and conditional random field," *Computational Intelligence*, vol. 38, no. 1, pp. 205–215, 2022.
- [12] M. Villamizar, O. Canévet, and J. M. Odobez, "Multi-scale sequential network for semantic text segmentation and localization," *Pattern Recognition Letters*, vol. 129, no. 1, pp. 63–69, 2020.
- [13] S. Gu and F. Zhang, "Applicable scene text detection based on semantic segmentation," *Journal of Physics: Conference Series*, vol. 1631, no. 1, Article ID 012080, 2020.
- [14] A. Ar, "Domain adaptation for part-of-speech tagging of Indonesian text using affix information - ScienceDirect," *Procedia Computer Science*, vol. 179, no. 1, pp. 640–647, 2021.
- [15] A. Chaudhary, A. Anastasopoulos, Z. Sheikh, and G. Neubig, "Reducing confusion in active learning for part-of-speech tagging," *Transactions of the Association for Computational Linguistics*, vol. 9, no. 1, pp. 1–16, 2021.

- [16] A. Rkm and A. Rb, "Integration of morphological features and contextual weightage using monotonic chunk attention for part of speech tagging," *Journal of King Saud University-Computer and Information Sciences*, vol. 27, no. 1, pp. 1–10, 2021.
- [17] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: probabilistic models for segmenting and labeling sequence data," *Proceedings of the IEEE*, vol. 11, 2001.
- [18] Y. Liang, X. Zhao, A. J. X. Guo, and F. Zhu, "Hyperspectral image classification with deep metric learning and conditional random field," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 6, pp. 1042–1046, 2020.
- [19] S. T. Cheng, C. W. Hsu, G. J. Horng, and S. Y. Chen, "Across-camera object tracking using a conditional random field model," *The Journal of Supercomputing*, vol. 77, no. 12, Article ID 14252, 2021.
- [20] 裴 Pei Liang, 刘 Liu Yang, and 高 Gao Lin, "Cloud detection of ZY-3 remote sensing images based on fully convolutional neural network and conditional random field," *Laser & Optoelectronics Progress*, vol. 56, no. 10, Article ID 102802, 2019.
- [21] J. Zheng and R. Ma, "Analysis of enterprise human resources demand forecast model based on SOM neural network," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6596548, 10 pages, 2021.
- [22] X. Chen, H. H. Wang, and B. Tian, "Visualization model of big data based on self-organizing feature map neural network and graphic theory for smart cities," *Cluster Computing*, vol. 22, no. S6, Article ID 13293, 2019.
- [23] H. Zhou and K. Yu, "A novel wireless sensor network data aggregation algorithm based on self-organizing feature mapping neural network," *Ingénierie des Systèmes d'Information*, vol. 24, no. 1, pp. 119–123, 2019.
- [24] K. J. Devi, N. H. Singh, and K. Thongam, "Automatic speaker recognition from speech signals using self organizing feature map and hybrid neural network," *Microprocessors and Microsystems*, vol. 79, no. 4, Article ID 103264, 2020.
- [25] D. Qiu, H. Xu, D. Luo et al., "A rainwater control optimization design approach for airports based on a self-organizing feature map neural network model," *PLoS One*, vol. 15, no. 1, Article ID e0227901, 2020.

Research Article

A Framework for Interactive Medical Image Segmentation Using Optimized Swarm Intelligence with Convolutional Neural Networks

Chetna Kaushal ¹, Md Khairul Islam ², Sara A. Althubiti ³, Fayadh Alenezi ⁴,
and Romany F. Mansour ⁵

¹Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

²Department of Information Communication Technology, Islamic University, Kushtia, Bangladesh

³Department of Computer Science, College of Computer and Information Sciences, Majmaah University, Al-Majmaah 11952, Saudi Arabia

⁴Department of Electrical Engineering, College of Engineering, Jouf University, Saudi Arabia

⁵Department of Mathematics, Faculty of Science, New Valley University, El-Kharga 72511, Egypt

Correspondence should be addressed to Md Khairul Islam; mdkito51@gmail.com

Received 3 June 2022; Revised 9 July 2022; Accepted 27 July 2022; Published 24 August 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Chetna Kaushal 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.

Recent improvements in current technology have had a significant impact on a wide range of image processing applications, including medical imaging. Classification, detection, and segmentation are all important aspects of medical imaging technology. An enormous need exists for the segmentation of diagnostic images, which can be applied to a wide variety of medical research applications. It is important to develop an effective segmentation technique based on deep learning algorithms for optimal identification of regions of interest and rapid segmentation. To cover this gap, a pipeline for image segmentation using traditional Convolutional Neural Network (CNN) as well as introduced Swarm Intelligence (SI) for optimal identification of the desired area has been proposed. Fuzzy C-means (FCM), K-means, and improvisation of FCM with Particle Swarm Optimization (PSO), improvisation of K-means with PSO, improvisation of FCM with CNN, and improvisation of K-means with CNN are the six modules examined and evaluated. Experiments are carried out on various types of images such as Magnetic Resonance Imaging (MRI) for brain data analysis, dermoscopic for skin, microscopic for blood leukemia, and computed tomography (CT) scan images for lungs. After combining all of the datasets, we have constructed five subsets of data, each of which had a different number of images: 50, 100, 500, 1000, and 2000. Each of the models was executed and trained on the selected subset of the datasets. From the experimental analysis, it is observed that the performance of K-means with CNN is better than others and achieved 96.45% segmentation accuracy with an average time of 9.09 seconds.

1. Introduction

To advance the efficiency and accuracy of the medical diagnostic system, especially those that are distributed in complex areas (e.g., brain, skin, lung, and blood cancer classification), several live line diagnostic models (Bengio et al. [1]; Tiwari et al. [2]; Bhatt et al. [3]) work with image processing. The effectiveness of the detection and accuracy of the multidisciplinary medical data diagnostic system depends

largely on the quality of the included images captured by a few techniques (Cetin et al. [4]) such as Magnetic Resonance Imaging (MRI) for brain data analysis, dermoscopic for skin, microscopic for blood leukemia, computed tomography (CT) scan images for lungs, etc. However, due to the uncontrollable lighting conditions and lots of noise availability during capturing, the illumination distributed on the surface of the medical images remains uneven, especially when backlight or fixed lighting conditions affect the diagnostic model

(Bonabeau et al. [5]; Banks et al. [6]; Siva Raja and Rani [7]). It also leads to a comparison of the global low and local image and weak data in the black region and each aspect of the image plays a significant function in the examination of medical data. Therefore, preliminary processing is an important part of medical image classification because it plays an important role in computer-assisted medical diagnostic programs in different systems. Due to the position variability of the targeted regions, traditional hybrid segmentation technique such as Fuzzy Competitive Learning based Counter Propagation Network (FCPN) still works better than soft computing techniques [8]. Moreover, image classification is considered the most important process for medical imaging as it extracts the region of interest (ROI) from various data by semiautomatic or automatic process (Kennedy and Eberhart [9]). It classifies the image in areas based on a specific definition, such as the segmentation of damaged body parts or tissues in the local medical diagnostic system and the acquisition of boundaries and classification. There are some samples of medical image segmentations shown in Figure 1.

For medical experts and researchers alike, performing good medical image segmentation is a difficult undertaking (Karaboga [10]). Many researchers, on the other hand, have previously attempted to create an effective algorithm for medical image segmentation in order to aid in the identification of various disorders and diseases. Therefore, in this research work, we present a comparative medical image segmentation framework using swarm intelligence with Convolutional Neural Networks (CNN). Mainly, three different medical image segmentation mechanisms are considered and studied, namely, Traditional Segmentation (Fuzzy C-means (FCM) and K-means), Swarm Intelligence (Particle Swarm Optimization (PSO)), and CNN-based segmentation with traditional approaches, as shown in Figure 2 with subclasses of architectures.

Used traditional segmentation techniques are the type of unsupervised machine learning and useful to find out groups or different patterns in medical data. In general terms, it's an unsupervised activity that divides unstructured data into several groups based on their similarity and dissimilarity (Yang, 2009 [11]). The motivation behind the proposed scenario is given in further sections.

1.1. Motivation. Unsupervised clustering-based medical image segmentation is a default method that aims to collect a set of objects or pixels into subsets or collections by the background and front of the image. The goal is to create clusters or parts that fit inside but are very different from each other. In simple terms, pixels in the same category should be as similar as possible, and objects in the same category should be very different from those in another cluster. Some challenging factors given that gave us the motivation are as follows:

- (i) The number of research articles available was large but lack of appropriate comparisons of Traditional medical image Segmentation, SI, and CNN-based segmentation.

- (ii) Medical data or image segmentation is a challenging task and still lots of improvements are needed to develop a better diagnosis system.
- (iii) Existing CNN-based models need a lengthy system-training period.
- (iv) Suffering from the over-fitting problems and need to solve such kind of problem regarding the medical diagnosis system that helps to detect the diseases in early stage. The overfitting problem in deep learning usually occurs when the image count is small in the target.
- (v) The existing system had to be developed and updated in real-time scenarios.
- (vi) There are no studies that have established a single standard segmentation model for distinct picture types from various organs.

1.2. Contributions. Nowadays, medical image segmentation using clustering is a basic requirement for lots of purposes like abnormal region detection, automatic extraction, data organization, etc. In the segmentation of medical data, high-quality clustering techniques are critical. Thus, in this research, we proposed a comparative framework for medical image segmentation with SI as well as CNN technique and the main contributions are as follows:

- (i) To study the existing medical image segmentation approaches with different algorithms.
- (ii) Develop a novel pre-processing for medical images like image quality enactment, hair removal from dermoscopic images, and blast nucleus improvement for microscopic images.
- (iii) To segment medical images, FCM and K-means are used as unsupervised machine learning approaches with PSO as swarm intelligence and CNN as a deep learning mechanism. A novel fitness function is presented here that replaces pixels to increase segmentation quality.
- (iv) To validate the proposed framework, performance parameters such as Precision, Recall, F-measure, Accuracy, Error, Matthews's Correlation Coefficient (MCC), Dice coefficient (DC), Jaccard Coefficient (JC), and time being calculated and related with existing works.

This research article deals with a comparative study for medical image segmentation and the rest of the article is systematized into different sections. Section 2 illustrates the survey of related work, and the methodology of the proposed mechanism is described in Section 3. In Section 4, results and discussion based on the performance parameters are illustrated, and Section 5 concludes with recommendations for the future.

2. Literature Survey

Segmentation is widely used in various sectors such as split geographical regions, fruit from trees, flood for damage reports, recognition of traffic signs, and road collapses.

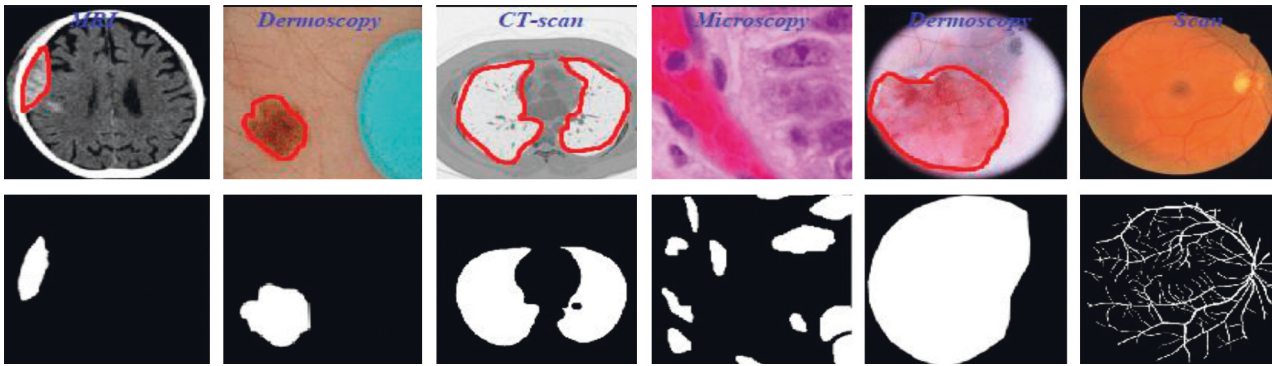


FIGURE 1: Medical image segmentation.

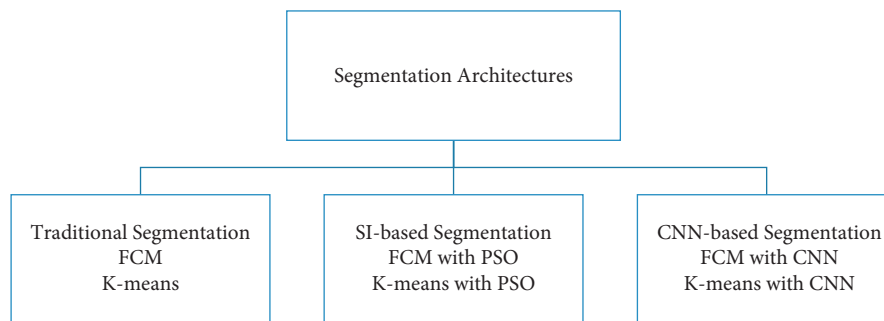


FIGURE 2: Proposed segmentation architectures.

Chouhan et al. [12] surveyed Computational Intelligence (CI) techniques to demonstrate the application of segmentation in the interdisciplinary research area. They discussed well-known CI techniques such as neural network, fuzzy logic, and genetic algorithms as well as they have also proved that CI-based approaches are cost-effective, time-saver, higher-efficient, and applicable in various engineering sectors [12, 13]. In 2021, researchers introduced an Internet of Things (IoT) device for automatic plant disease (galls) detection using the Fuzzy Based Function Network (FBFN) segmentation technique [14]. Similarly, a web-based tool was developed to identify mango leaf diseases such as Anthracnose using the RBF segmentation method. Thus, the segmentation techniques are universally implemented for different image-based detection techniques [15].

However, image segmentation is extensively used in human disease detection and diagnosis. For precise detection of the disease, initially, it requires identifying the region of interest from the captured images. In this study, a comprehensive description of the most important state-of-the-art medical image segmentation techniques is given. Here, we consider a mixed survey of segmentation for different types of medical data.

IMV-FCM, an enhanced multiscreen FCM clustering method, introduces a weighted adaptive learning technology to increase the flexibility of coordinating from diverse viewpoints. The algorithms might be able to learn from each view in an adaptive way that helps them better group brain tissue and deal with a noise like partial dimension distortions and grayscale that does not match up [16]. In 2018,

Karegowda et al. [17] conducted research on the segmentation of brain tumor regions from MRI data. The authors conducted a comparative examination of FCM, Adaptive Regularized Kernel-based FCM approaches, PSO, and K-means and concluded that using PSO as swarm intelligence is a useful step. The results of the experiments showed that PSO-based segmentation is more accurate than FCM, Adaptive Regularized Kernel FCM, and K-means [17]. Arun Kumar et al. [18] created an improved automated approach for segmenting brain tumor regions and identifying them using K-means for the same objective. The goal of the authors was to improve the imaging enhancement at the pre-processing stage for precise brain tumor prediction [18]. Chander et al. [19] developed a framework for the segmentation of MRI images using K-means with Support Vector Machine (SVM) as the machine learning approach, and the overall accuracy was increased over earlier work [19]. In [20], the authors have discussed various methods such as traditional segmentation (Threshold, Fuzzy Theory, Region and edge detection), machine learning approach (KNN, Random Forest, SVM, Dictionary learning), and deep learning methods (CNN, FCM, Encoder/decoder). Although their analysis depicts that deep learning-based techniques such as FCM are superior to other traditional methods, they also include a supervised method that demands for manual labeling which required domain-specific knowledge [20].

In one of our previous studies, we have developed an IoT-based data collection system for skin lesions where we classified various skin lesions using deep learning-based

ensemble algorithms [21]. For skin lesion segmentation, Yuan et al. [22] used the notion of Deep Fully Convolutional Neural Networks (DFCNN) with Jaccard distance. They employed the 19-DFCNN layer for self-training and the function of the new loss based on the Jaccard scale created by the researchers to re-measure using cross entropy to distinguish the lesion from the skin lesions. The findings of the studies imply that the upgraded classification approach outperforms conventional state algorithms, but that it requires more pre- and post-processing stages for greater accuracy [22]. Using CNN, Xie et al. [23] devised a reliable approach for extracting skin lesion bounds in the existence of distortions in digital images. Due to the use of a basic segmentation strategy, detecting the boundary of a skin lesion zone is slow, but this can be solved by utilizing a semantic segmentation technique [23]. In 2022, the authors trained a feature adaptive transformers network (FAT-Net) and managed to handle blurred boundary issues associated with lesions image. Yet FAT-Net may effectively extract local features and global true label whereas CNN are not capable of learning global true labels sufficiently [24]. Similarly, a neural network-based Multi-scale Residual Encoding and Decoding network (Ms RED) is used to handle blurred boundaries [25]. Thapar et al. [26] employed a segmentation framework using swarm intelligence with Grasshopper Optimization Algorithm (GOA) for feature extraction and successfully obtained 98.42% classification accuracy. Nevertheless, they only trained the model on three skin lesions images [26].

The existence of the nucleus in blood cells is used in determining Leukemia. In 2021, Daud et al. [27] used conventional algorithms such as watershed distance transform and Sobel edge detection algorithm for segmenting nuclei from microscopic images [27]. In another research, authors deployed a Global Local Entropy Histogram Equalization (GLEHE) based segmentation technique to identify Leukemia in blood cells [28]. Dhal et al. [29] provided a method for segmenting blood images for leukemia using the Stochastic Fractal Search (SFS) algorithm, which provides non-false positive segmented results. For image segmentation, the notion of K -means-based clustering is studied. The proposed scheme was compared to a previous clustering method, and the findings showed that the system's performance was better in terms of efficiency, computational burden, and quality attributes [29].

Senthil Kumar et al. [30] used five algorithms to extract a plant region from very small lung images, including PSO, inertia-weighted PSO, guaranteed convergence PSO (GCPSO), K -means, and K -median. The flexible median filter outperformed the central filters, intermediate variables, and standard pre-processing stage, proving that it is best suited for medical CT imaging. In addition, employing the changing histogram balance improves the image brightness. Four algorithms are used to determine the quality of pre-processed images with improved quality. GCPSO has a high accuracy of 95.89 percent when visual results were confirmed with 20 lung sample images using MATLAB [30]. In 2021, van De Worp et al. [31] introduced deep learning-based two-step U-Net architecture for lung cancer

segmentation from CT images. Although they performed the task only on 60 CT images [31]. The authors in [32] deployed 2-D Discrete Wavelet Transform (DWT) on the "LOTUS dataset" of lung tumor (CT images) and achieved a dice coefficient of 0.8472.

We give a quick summary of the literature review and the following aspects highlighted as limitations based on the preceding analysis:

- (i) The primary flaw with present clustering-based segmentation methods is that the foreground and background are overlapping.
- (ii) Bio-inspired algorithms are commonly utilized in optimization-based techniques, which require longer to complete the segmentation process due to the unknown high number of clusters (Kaushal et al. [33]).
- (iii) Because of the image quality, enormous segmentation tasks have suffered from difficulties in segmentation of complex images in cases of computed tomography scans, MRI, microscopic, and dermoscopic image modalities (Kaushal et al. [34]). It is necessary to focus on quality improvisation.
- (iv) Researchers encounter a pixel-mixing difficulty due to frequent pixel value changes in the region.

In this study, we are going to make it an intuitive and easy-to-understand framework for medical image segmentation.

3. Methodology

This section of the research article includes the procedural and working steps of the proposed model for Medical Image Segmentation using the Traditional Segmentation, SI, and CNN mechanisms. We focused on introducing a modified medical image segmentation approach using CNN as a deep learning and three distinct proposed architectures, which are as given in further sections.

3.1. Traditional Segmentation. In this phase, we evaluate the two clustering-based segmentation approaches such as FCM and K -means because it has many applications in medical research.

3.1.1. FCM-Based Segmentation. This scenario presents medical image segmentation using the concept of FCM as an unsupervised process. After segmentation, two parts of an image are formed known as the background and foreground part where the foreground is the ROI of any medical data such as MRI, microscopic, dermoscopic, CT-scan, etc. Here, we use some common pre-processing steps in the entire six scenarios of the proposed framework for comparative analysis, first is the color conversion using

$$G_{\text{Image}} = 0.3 \times +0.59 \times +0.11 \times, \quad (1)$$

where G_{Image} is the grey level image that obtained after the color conversion from the color images (RGB \rightarrow Red, Green, and Blue plane). After that, grey level mapping is

initiated on the clipped region of the image for quality enhancement using the given equation:

$$X_{\text{AVERAGE}} = \frac{X_{(\text{region-axis})} \times X_{(\text{region-axis})}}{G_{\text{Image}}}. \quad (2)$$

Equation (2) defines the average number of pixels in the medical image. Where $X_{(\text{region-axis})}$ is the total number of medical image pixels in a clipped region (X_{CLIP}). The clip limit (X_{CL}) of medical image enhancement is calculated using equation 3 then we apply the image enhancement of the further processing using the written Algorithm 1:

After medical image enhancement in pre-processing phase, we move toward the segmentation using the FCM and the FCM algorithm written as below in Algorithm 2.

The concept of FCM is dependent on the idea of consistently acquiring cluster centers by adjusting their positions using mean values that are given in equation 4 and allows clusters that are more flexible by introducing the possibility of partial memberships. The error function of FCM is written in the following equation:

$$\text{Error}_{\text{FCM}} = \sum_{i=1}^m \sum_{j=1}^n \mu_{ij}^k X_i^{(j)} - C_j^2, \quad (3)$$

where fuzzy membership is denoted by μ_{ij} of X_i (image's pixel) and the cluster identified by its center C_j , and here, k is a constant that defines the fuzziness of the resulting partitions.

$$\mu_{ij} = \frac{1}{\sum_{m=1}^C X_j - C_j / X_j - C_m^{2/(k-1)}}. \quad (4)$$

The steps involved in the algorithm of FCM-based segmentation are

FCM 1: recruit C_i as the cluster centers and Iteration $N=0$

FCM 2: call FCM membership functions μ_{ij} according to equation 7

FCM 3: let $N=N+1$ and assign new C_i as new centers

FCM 4: until the best convergence is not found, repeat steps 2 to 3.

Using this algorithm, we segment the ROI from the medical images and after segmentation of medical images; the obtained segmented result with original images shown in Figure 3.

3.1.2. K-Means Based Segmentation. This is the second scenario and we used K -means as a segmentation technique instead of FCM because K -means helps to provide better segmentation results as compared to the FCM that is shown in Figure 4. By utilizing the concept of K -means as a medical image segmentation technique, appropriate ROI from the medical images could be segmented but also K -means faced mix-up issues, and the algorithm of K -means is written in Algorithm 3.

Based on the above written K -means algorithm in the ASBT system, we obtained better-segmented result as compared to the FCM-based ASBT system, and the results with the original MRI image are shown in Figure 4.

3.2. SI-Based Segmentation. In this scenario, we the concept of PSO as a SI approach because it is the most well-known optimization technique that helps to optimize the pixel-mixing problem faced by the FCM. Here, we present two different hybrid mechanisms named FCM and K -means with PSO for medical image segmentation.

3.2.1. FCM with PSO-Based Segmentation. In this scenario, we utilize the concept of PSO along with the FCM as a medical image hybrid segmentation. PSO is a powerful meta-heuristic technique that is favored by birds, mammals, and other insects that live or move in swarms. The best example of PSO inspiration is a flock of birds or a school of tiny fish that helps to reduce the pixel-mixing problem faced by the FCM during medical image segmentation by having very similar neighborhood pixel values. This hybrid segmentation method is based on the participation of individuals of each particle from many fields that are participating in the searching mechanism of a threshold value to minimize the mixed pixels. To solve a medical image segmentation problem, each particle modifies its threshold value based on its own and its neighbors' experiences. Formally, each PSO particle P_I has a position $P_I(t)$ at the time t instances in the search space, which change at time $t+1$ by a velocity $V_I(t)$. In the PSO algorithm, $V_I(t)$ velocity is influenced by the best position $V_{\text{BEST}}(t)$ visited by itself and $P_{\text{ALL}}(t)$ the best position visited by all particles (we termed it "global best"). Each particle's position is determined by a unique fitness function (Fit (Fun)), which is dependent on the segmentation issue and space dimension D .

$$\begin{aligned} P_I(t) &= P_{I1}, P_{I2}, P_{I3}, \dots, P_{ID}, \\ V_I(t) &= V_{I1}, V_{I2}, V_{I3}, \dots, V_{ID}, \\ V_{\text{BEST}}(t) &= V_{\text{BEST}1}, V_{\text{BEST}2}, V_{\text{BEST}3}, \dots, V_{\text{BEST}D}, \\ P_{\text{ALL}}(t) &= P_{\text{ALL}1}, P_{\text{ALL}2}, P_{\text{ALL}3}, \dots, P_{\text{ALL}D}. \end{aligned} \quad (5)$$

Kennedy and Eberhart, (1995) [9] established the PSO algorithm as an evolutionary image segmentation technique, and the algorithm of FCM with PSO segmentation is written below in Algorithm 4.

Based on the above-written hybrid segmentation algorithm using FCM with PSO, we obtained better-segmented results as compared to the only FCM as well as K -means also and results with the original medical image shown in Figure 5.

3.2.2. K-Means with PSO-Based Segmentation. The concept of PSO along with the K -means clustering algorithm used as a medical image hybrid segmentation and the algorithm of K -means with PSO segmentation is written below in Algorithm 5.

```

Input:  $M \rightarrow$  Medical-Image
Output:  $EM \rightarrow$  Enhanced Medical Image
Start enhancement process
Set clip limit,  $X_{CL} = X_{CLIP} - X_{AVERAGE}$ 
  [Row, Col., and Plane] = Size ( $M$ )
If Plane > 1//Means image is color
   $CM = \text{RGB to Grey } (M)$ 
Else
   $CM = M$ //No need of conversion
End-If
For  $I = 1 > X_{CL}$ 
   $EM = \text{Intensity } (M (I), X_{CL})$ 
End-For
Return:  $EM$  as an enhanced medical image
End-Algorithm

```

ALGORITHM 1: Enhancement of medical images.

```

Input:  $EM \rightarrow$  Enhanced Medical Image
Output:  $BM$  and  $FM$  (ROI)  $\rightarrow$  Background and Foreground
Start FCM
  Define cluster number ( $G = 2$ )
  [Row, Col.] = Size ( $EM$ )
  Segregate  $G = G1 \& G2$  /Where  $G1$  for  $BM$  and  $G2$  for  $ROI$ 
  Rep =  $N$ //Define number of iterations for clustering
  While Rep  $\neq N$  (until max iteration not achieved)
    For  $P = 1 \rightarrow \text{Row}$ 
      For  $Q = 1 \rightarrow \text{Col}$ 
        If  $EM (P, Q) = G1$ 
           $BM (P, Q) = EM (P, Q)$ 
        Else if  $EM (m, n) = G2$ 
           $ROI (P, Q) = EM (P, Q)$ 
        End-If
      Adjust Centroid  $G$  using given Algorithm 2
       $G_{mn} =$ 
      Repeat and define FCM membership using given Algorithm 2
       $[G1, G2] = \sum_1^n (d_{Gm}^2 / d_{Gn}^2)^{1/m-1}$ 
    End-For
  End-For
End-While
Return:  $BM$  and  $ROI$ 
End-Algorithm

```

ALGORITHM 2: FCM-based Segmentation.

Based on the above-written hybrid segmentation algorithm using K-means with PSO, we obtained better-segmented results as compared to the FCM, K-means, and FCM with PSO also and results with the original medical image shown in Figure 6.

3.3. CNN-Based Segmentation. This is the third module of implementation where we used again two different scenarios that are described below.

3.3.1. FCM with CNN-Based Segmentation. In this scenario, we utilize the concept of CNN as a deep learning mechanism along with the FCM as a medical image hybrid segmentation. This hybrid method is currently used in most of the existing medical image segmentation research. First, we train the model using lots of already segmented images in terms of background and foreground images having 3 dimensions (RGB). Usually, segmented medical images, which are fed into the neural network, are reduced in data dimensions, reduce the

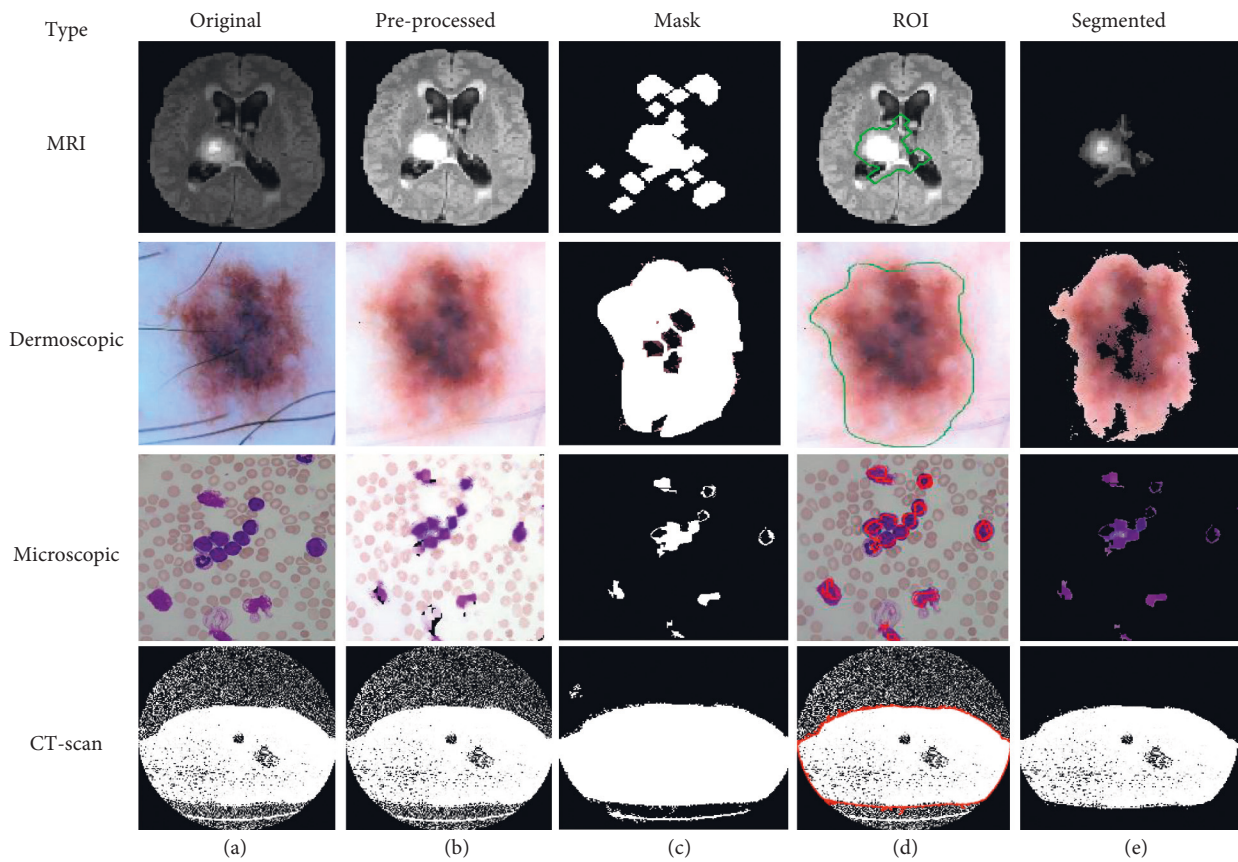


FIGURE 3: (a) Original. (b) Pre-processed. (c) Mask. (d) ROI. (e) Segmented image using FCM.

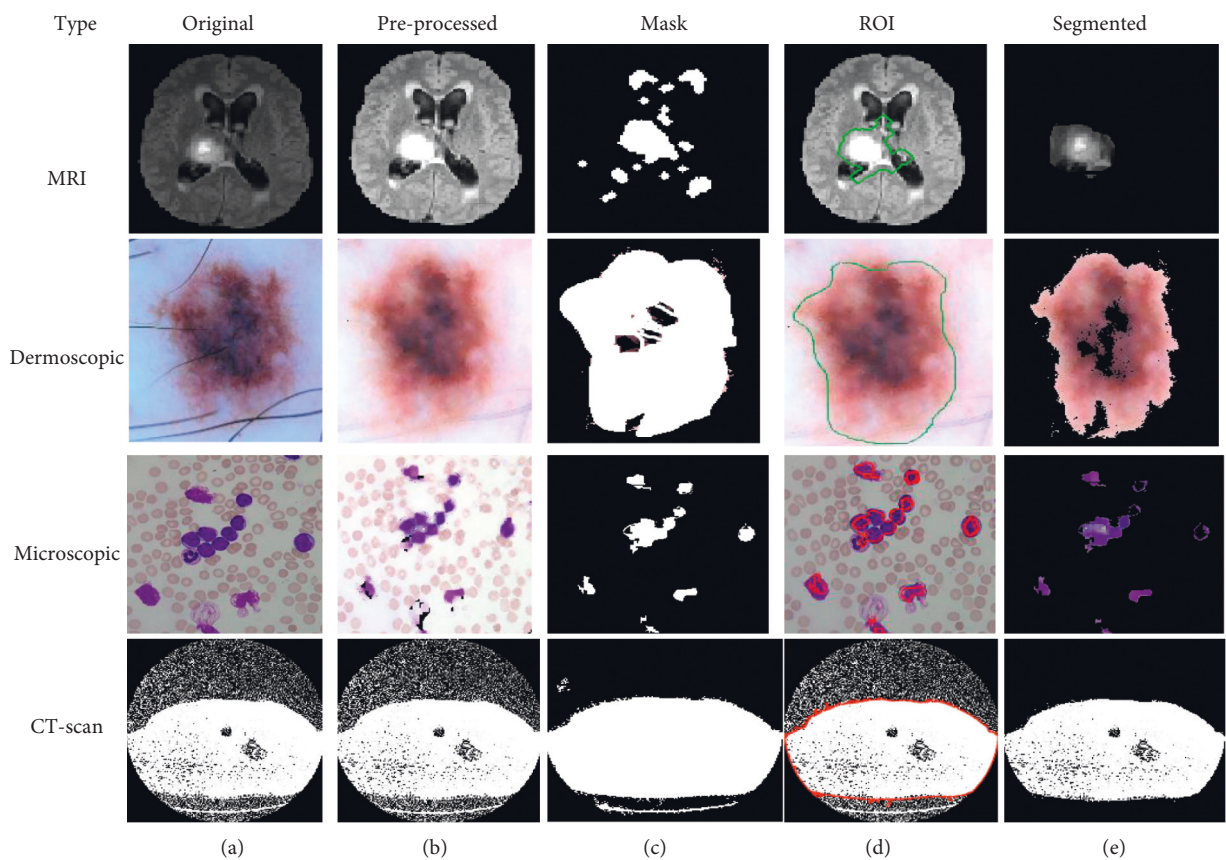


FIGURE 4: (a) Original. (b) Pre-processed. (c) Mask. (d) ROI. (e) Segmented image using *K*-means.


```

Input: EM → Enhanced Medical Image
Output: BM and FM (ROI) → Background and Foreground
Start K-means
Define cluster number ( $G = 2$ )
[Row, Col.] = Size (EM)
Segregate  $G = G1$  &  $G2$ //Where  $G1$  for BM and  $G2$  for ROI
Rep =  $N$ //Define number of iterations for clustering
While Rep  $\neq N$  (until max iteration not achieved)
For P = 1 → Row
  For Q = 1 → Col
    If EM (P, Q) == G1
      BM (P, Q) = EM (P, Q)
    Else if EM (m, n) == G2
      ROI (P, Q) = EM (P, Q)
    End-If
  Adjust Centroid C using their mean
   $G = \text{Average (BM, ROI) using the given Algorithm 3}$ 
   $G_{mn} = \sum_{m=1}^{\text{Row}} \sum_{n=1}^{\text{Col}} G1_{mn} + G2_{mn}/2$ 
  End-For
End-For
End-While
Return: BM and ROI
End-Algorithm

```

ALGORITHM 3: K-means based Segmentation.

```

Input: EM → Enhanced Medical Image
Output: BM and FM (ROI) → Background and Foreground
Start Hybridisation
Define cluster number ( $G = 2$ )
[Row, Col.] = Size (EM)
Segregate  $G = G1$  &  $G2$ //Where  $G1$  for BM and  $G2$  for ROI
Rep =  $N$ //Define number of iterations for clustering
While Rep  $\neq N$  (until max iteration not achieved)
For P = 1 → Row
  For Q = 1 → Col
    If EM (P, Q) == G1
      BM (P, Q) = EM (P, Q)
    Else if EM (m, n) == G2
      ROI (P, Q) = EM (P, Q)
    End-If
  Adjust Centroid G using given Algorithms 1 and 2
  End-For
End-For
End-While
  To optimize the ROI, here we used PSO algorithm and then initialize using the following parameters such as:
  (i) Iterations ( $T$ )
  (ii) Population-size ( $S$ )
  (iii) Lower-Limit (LB)
  (iv) Upper-Limit (UB)
  (v) Fitness function
  (vi) Number of selection ( $N$ )
  Calculate size in terms of  $T = \text{Row} \times \text{Col}$ .
  Fitness function for the parameter optimization of PSO:
  
$$\text{fit}(\text{fun}) = \begin{cases} 1 & \text{if pixel is less} \\ 0 & \text{otherwise} \end{cases}$$

  For I = 1 → T
   $fs = \text{EMRI}(I)$ 

```

ALGORITHM 4: Continued.

```


$$ft = \sum_{i=1}^{\text{Pixels}} \text{EMRI}(l) / \text{LengthofEMRIPixels}$$

fit(fun) = using Algorithm 4
Thresholdvalue = PSO(P, T, LB, UB, N, fit(fun))
End-For
Define optimization iterations, O-Rep = N
While O-Rep ≠ N (until max iteration not achieved)
Thr = Thresholdvalue
Mask = Binary(ROI, Thr)
ROI Boundaries = Boundary(Mask)
For k = 1 → D
ROI = EM × ROI Boundaries
End-For
End-While
Return: BM and ROI as an improved background and foreground
End-Algorithm
    
```

ALGORITHM 4: FCM with PSO based Segmentation.

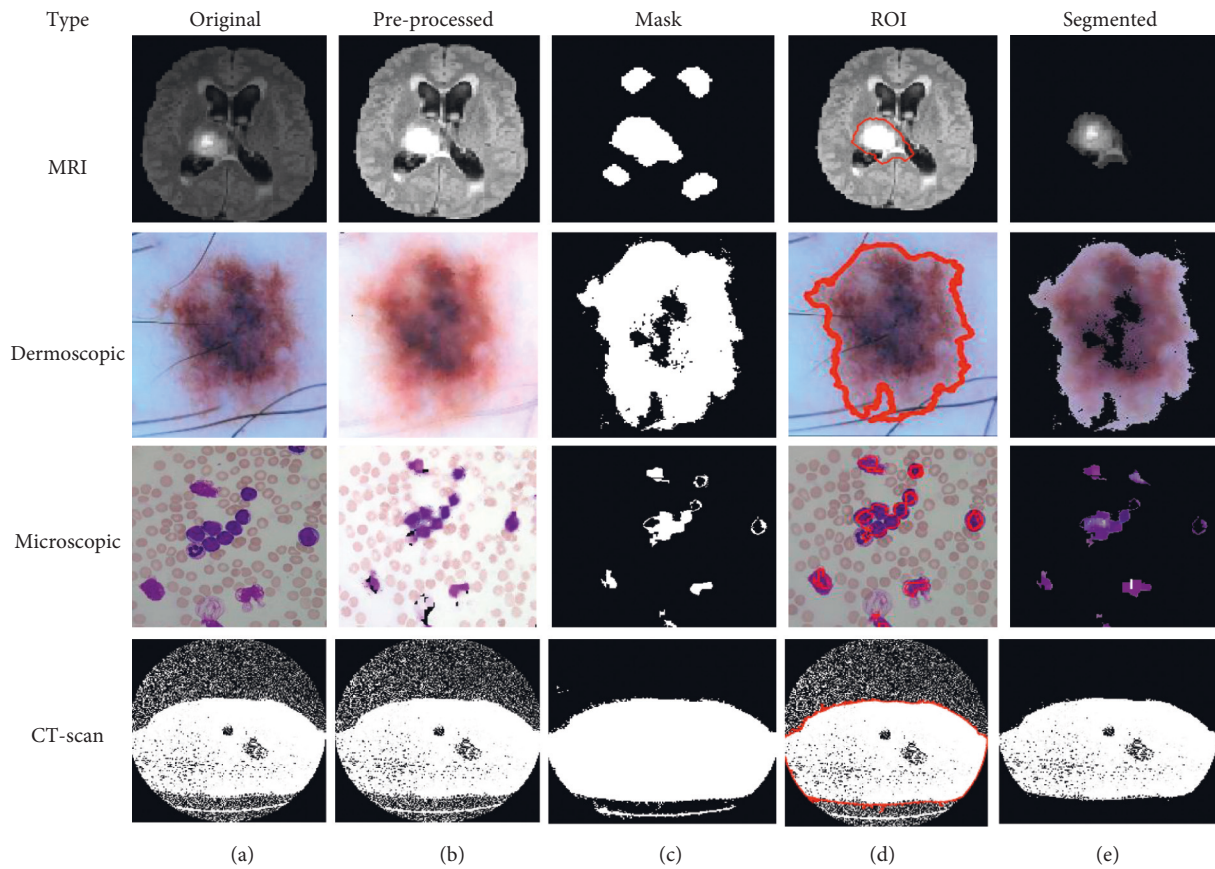


FIGURE 5: (a) Original. (b) Pre-processed. (c) Mask. (d) ROI. (e) Segmented image using FCM with PSO.

system processing time as well as complexity and help to reduce the over-fitting problems and hybrid CNN mechanism shown in Figure 7.

Figure 7 illustrated the process of medical image segmentation using the hybridization of FCM with CNN and the algorithm of K-means with PSO segmentation written below in Algorithm 6.

We obtained better-segmented results for the proposed hybrid mechanism of FCM with CNN as compared to the

FCM, K-means, and improvisation in FCM with PSO, improvisation in K-means with PSO and results with the original medical image shown in Figure 8.

3.3.2. *K-Means with CNN-Based Segmentation.* The concept of CNN as a deep learning mechanism along with the K-means used in this scenario and similar to algorithm 6, a hybrid CNN mechanism with K-means is shown in Figure 9.

Input: EM \rightarrow Enhanced Medical Image
Output: BM and FM (ROI) \rightarrow Background and Foreground
Start Hybridisation
 Define cluster number ($G = 2$)
 [Row, Col.] = Size (EM)
 Segregate $G = G1 \& G2$ // Where $G1$ for BM and $G2$ for ROI
 Rep = N // Define number of iterations for clustering
While Rep $\neq N$ (until max iteration not achieved)
 For P = 1 \rightarrow Row
 For Q = 1 \rightarrow Col
 If EM (P, Q) = $G1$
 BM (P, Q) = EM (P, Q)
 Else if EM (m, n) = $G2$
 ROI (P, Q) = EM (P, Q)
 End-If
 Adjust Centroid C using their mean
 $G = \text{Average (BM, ROI)}$ using the given Algorithm 5
 $G_{mm} = \sum_{m=1}^{\text{Row}} \sum_{n=1}^{\text{Col}} G1_{mn} + G2_{mn}/2$
 End-For
 End-For
End-While
 To optimize the ROI, here we used PSO algorithm and then initialize using following parameters such as:
 (i) Iterations (T)
 (ii) Population-size (S)
 (iii) Lower-Limit (LB)
 (iv) Upper-Limit (UB)
 (v) Fitness function
 (vi) Number of selection (N)
 Calculate size in terms of $T = \text{Row} \times \text{Col}$.
Fitness function using Algorithm 4 for the parameter optimization of PSO
For I = 1 \rightarrow T
 $fs = \text{EMRI}(I)$
 $ft = \sum_{i=1}^{\text{Pixels}} \text{EMRI}(I) / \text{LengthofEMRIPixels}$
 fit (fun) = using Algorithm 4
 Threshold_{value} = PSO ($P, T, LB, UB, N, \text{fit}(\text{fun})$)
End-For
 Define optimization iterations, O-Rep = N
While O-Rep $\neq N$ (until max iteration not achieved)
 Thr = Threshold_{value}
 Mask = Binary (ROI, Thr)
 ROI Boundaries = Boundary (Mask)
For k = 1 \rightarrow D
 ROI = EM \times ROI Boundaries
End-For
End-While
Return: BM and ROI as an improved background and foreground of medical image
End-Algorithm

ALGORITHM 5: K-means with PSO based Segmentation.

Figure 9 illustrated the process of medical image segmentation using hybridization of FCM with CNN and the algorithm of K -means with PSO segmentation written below in Algorithm 7.

In comparison to FCM, K -means, FCM with PSO, CNN, and K -means with PSO, we achieved better-segmented results for the proposed hybrid mechanism of K -means with CNN, and results with the original medical image are displayed in Figure 10.

Finally, performance parameters for different types of datasets are calculated and compared using a comparison

framework simulation in terms of Precision, Recall, F -measure, Accuracy, Error, MCC, DC, JC, and time.

3.4. Collected Dataset

3.4.1. *Brain Tumor Segmentation (BraTS) Dataset.* The sample images of the BraTS dataset are shown in Figure 11, it is a standard dataset obtained from “<https://www.med.upenn.edu/sbia/brats2018/data.html>” having MRI images [35]. For the simulation of the model, 50 DICOM files were

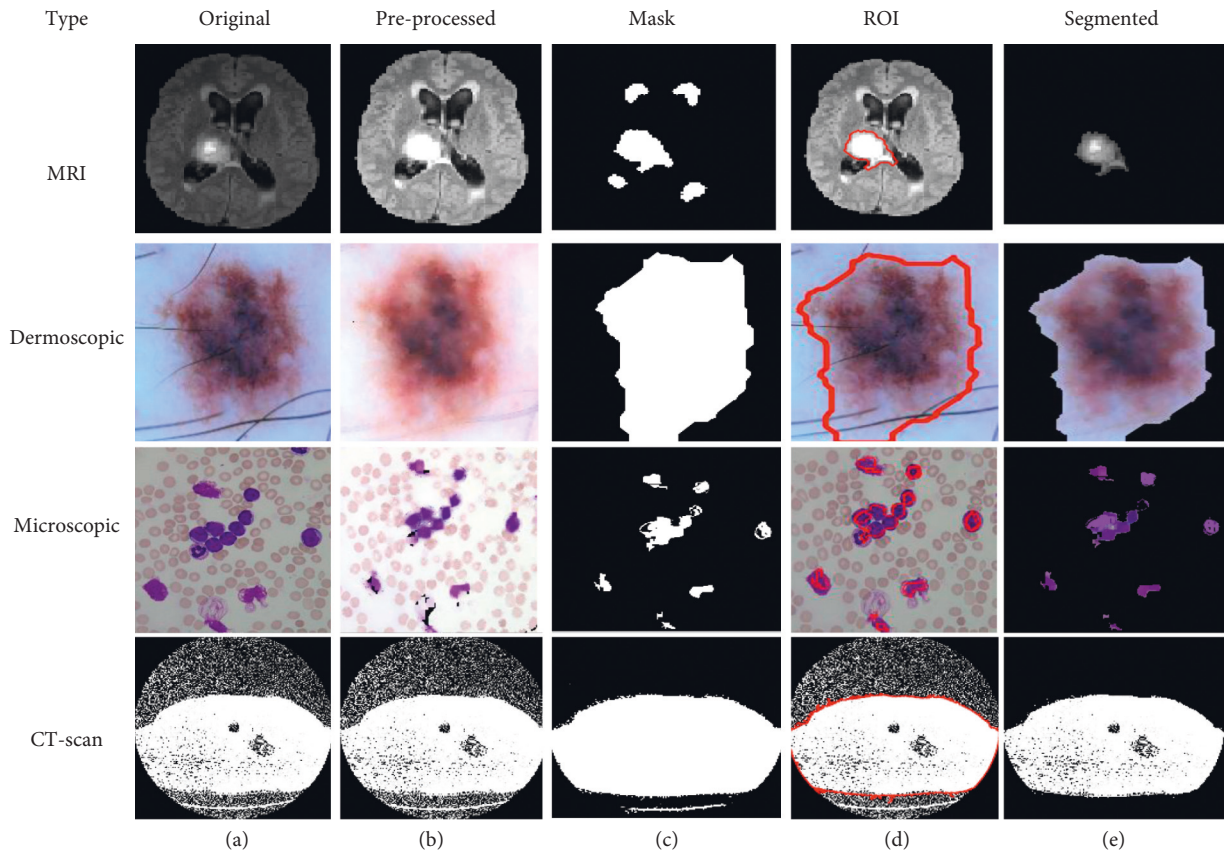


FIGURE 6: (a) Original. (b) Pre-processed. (c) Mask. (d) ROI. (e) Segmented image using K-means with PSO.

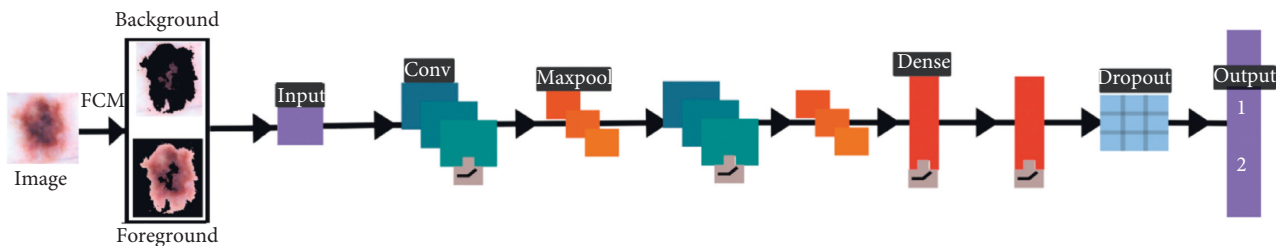


FIGURE 7: FCM with CNN for segmentation.

converted into JPG format that is representing multi-frame superimposed brain images.

3.4.2. Acute Lymphoblastic Leukaemia Image Database (ALL-IDB) Dataset. The dataset contains 2008 images that were collected in September 2005 in the Image Processing Department of Computer Science-Università degli Studi di Milano” [36]. The ALL-IDB dataset of microscopic images is freely available for scientific research purposes from “https://homes.di.unimi.it/scotti/all/” and the sample of ALL-IDB dataset images is shown in Figure 12.

The used dataset contains approximately 39,000 blood counts, and oncologists labeled the lymphocytes. We resize the original microscopic images of blood samples into a size of 256×256 and a total of 2000 images were used in this research work.

3.4.3. ISIC-2018 Dataset. It contains the human lesion analysis toward melanoma detection and the dataset is in the form of dermoscopic images. The dataset is available from <https://challenge2018.isic-archive.com/task1/training/> [37]. To capture images, the dermoscopic process is used which is an imaging technique to eliminate the surface reflection of human skin. It provides improved diagnostic accuracy and the sample of the ISIC-2018 dataset is shown in Figure 13.

3.4.4. CT-Scan Dataset. The database currently consists of an image set of 50 lung CT scans for research purposes which is publicly available from “http://www.via.cornell.edu/lungdb.html” [38]. A sample of dataset CT-scan images is shown in Figure 14.

```

Input: EM  $\rightarrow$  Enhanced Medical Image
Output: BM and FM (ROI)  $\rightarrow$  Background and Foreground
Start Hybridisation
Apply FCM and create BM and ROI
To optimize the ROI, here we used CNN and firstly we need to initialize using Epochs, Iterations, and Neurons (N) etc.
For I = 1  $\rightarrow$  T = Row  $\times$  Col.
If EM (I) belongs to BM
    Group (1) = Background
Else//EM belongs to ROI
    Group (2) = Foreground
End-If
End-For
Call pattern net CNN for training of system using BM and ROI data
FCM-CNN = Pattern-net (Neurons)
FCM-CNN = Train (FCM-CNN, EM, Group)
ROI = Sim (FCM-CNN, ROI data)
If ROI (Pixel) = ROI
    ROI = EM
Else
    BM = EM
End-If
Return: BM and ROI as an improved segmented background and foreground of medical image
End-Algorithm

```

ALGORITHM 6: FCM with CNN based Segmentation.

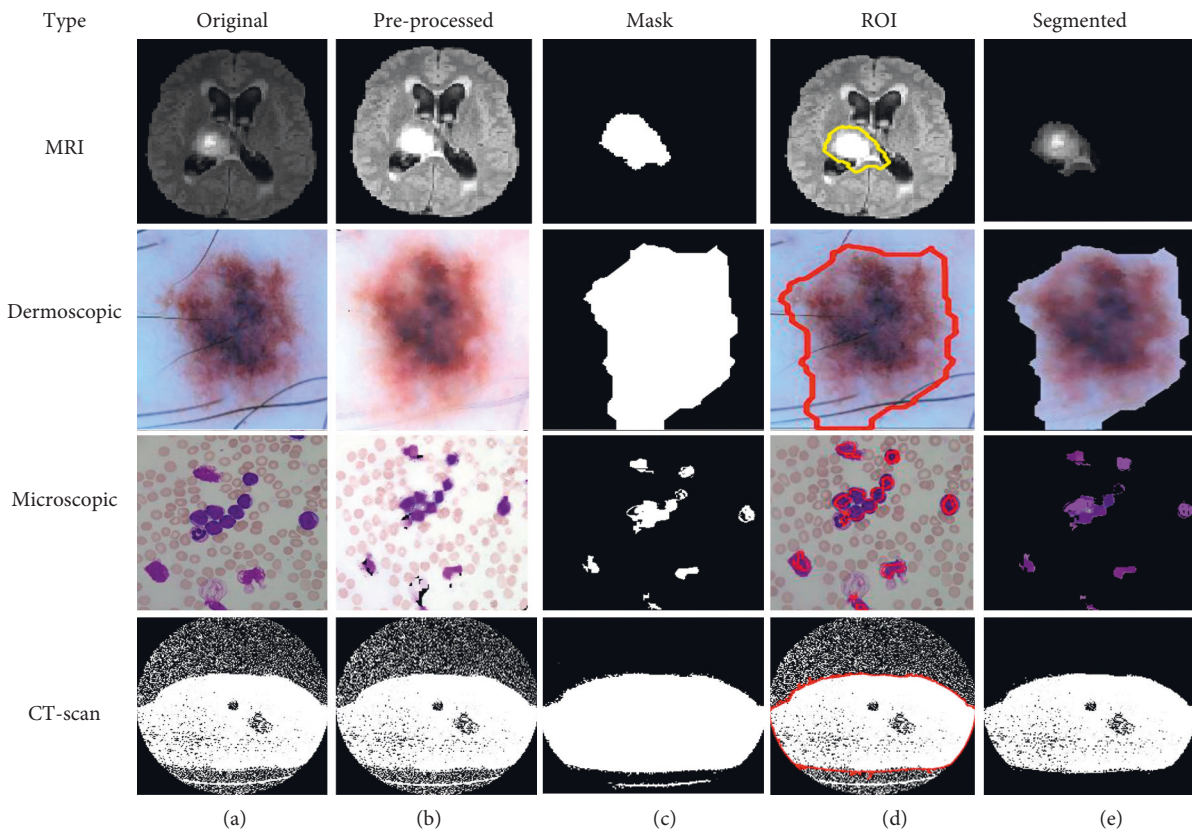


FIGURE 8: (a) Original. (b) Pre-processed. (c) Mask. (d) ROI. (e) Segmented image using FCM with CNN.

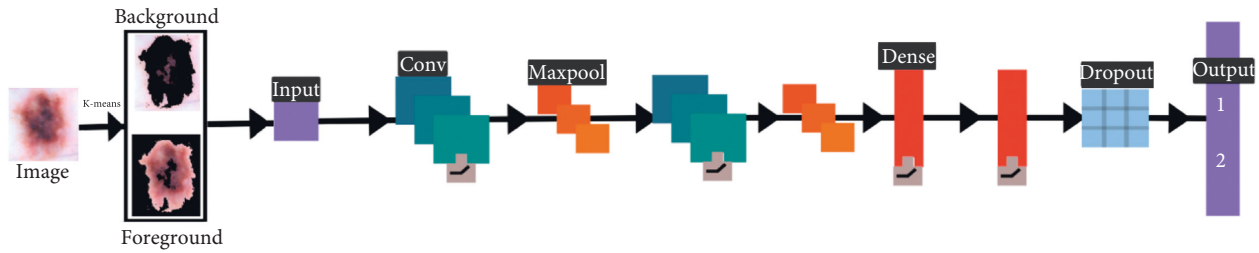


FIGURE 9: K-means with CNN for segmentation.

```

Input: EM → Enhanced Medical Image
Output: BM and FM (ROI) → Background and Foreground
Start Hybridisation
Apply K-means and create BM and ROI
To optimize the ROI, here we used CNN, and first, we need to initialize using Epochs, Iterations, Neurons (N) etc.
For I = 1 → T = Row × Col.
If EM (I) belongs to BM
    Group (1) = Background
Else//EM belongs to ROI
    Group (2) = Foreground
End-If
End-For
Call pattern net CNN for training of system using BM and ROI data
FCM-CNN = Pattern-net (Neurons)
FCM-CNN = Train (FCM-CNN, EM, Group)
ROI = Sim (FCM-CNN, ROI data)
If ROI (Pixel) = ROI
    ROI = EM
Else
    BM = EM
End-If
Return: BM and ROI as an improved segmented background and foreground of medical image
End-Algorithm

```

ALGORITHM 7: K-means with CNN based Segmentation.

After executing the methodology's outlined steps, the performance has been evaluated in terms of several parameters, as discussed in the result and discussion section.

3.5. Evaluation Metric. In this part, we have outlined the assessment measures used to verify the effectiveness of the suggested techniques. First of all, we have observed quantitative metrics such as Accuracy. In most cases, the efficiency of a model is measured in terms of its accuracy. However, in medical image segmentation, the model's accuracy is insufficient to provide a precise understanding of the model. Therefore, there are several additional measures, such as precision, recall, Error, and *F1* score, to assess segmentation quality. In order to analyze and comprehend the ability of the models, we have made use of each of these measures.

Moreover, we have considered similarity metrics such as Matthews's Correlation Coefficient (MCC), Dice coefficient (DC), and Jaccard Coefficient (JC). Each similar metric has a few special characteristics to evaluate the true performance of the selected segmentation

techniques. If all of the probabilistic methods, including true positives, true negatives, false negatives, and false positives, provide a high score, then the MCC algorithm will generate a higher score [39]. Similarly, DC deals with the missing data in image segmentation-related problems [40]. Both quantitative and similar metrics are considered in our study, which provides more robust comparisons and preferences of the specified segmentation techniques [41].

4. Results and Discussion

In this research work, we proposed a comparative framework for the medical image segmentation from various types of images such as MRI, Dermoscopic, Microscopic, and CT-scan images using the six different scenarios such as FCM, *K*-means, and improvisation of FCM using PSO, improvisation of *K*-means with PSO, improvisation of FCM with CNN and improvisation of *K*-means with CNN. Simulation results of the offered scenario are shown in Table 1 based on the quantities parameters.

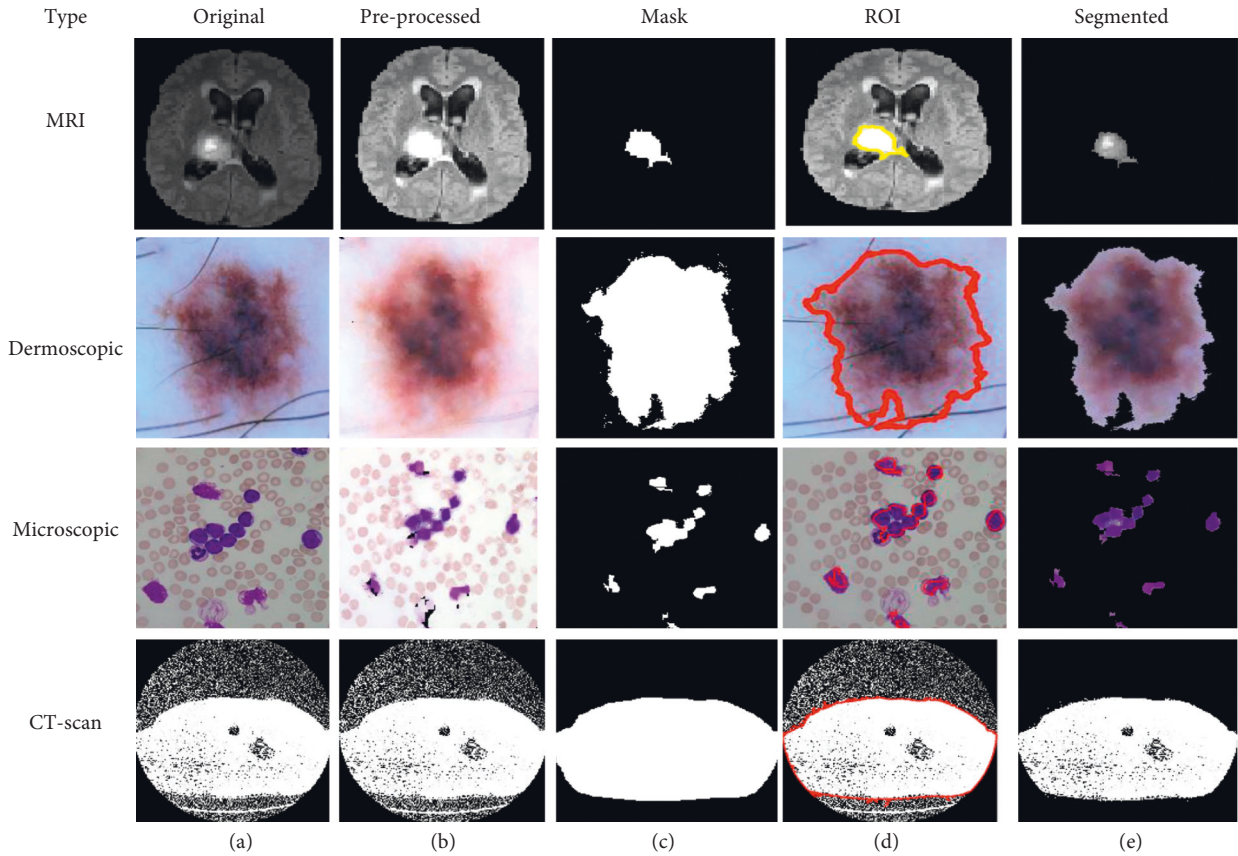


FIGURE 10: (a) Original. (b) Pre-processed. (c) Mask. (d) ROI. (e) Segmented image using *K*-means with CNN.

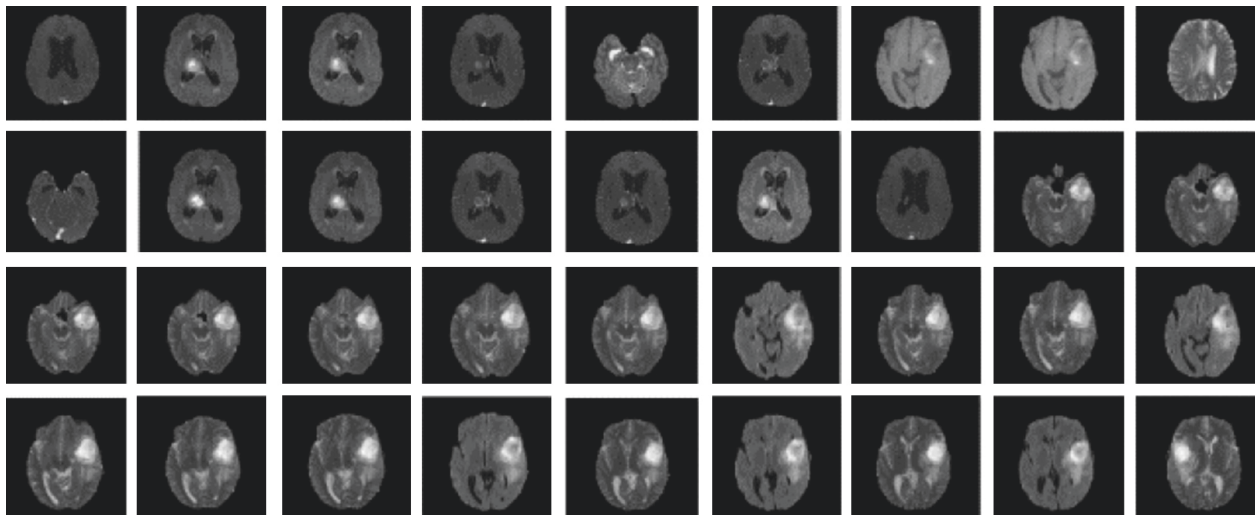


FIGURE 11: BraTS dataset.

There are five separate sets of data that included a varied number of images considering 50, 100, 500, 1000, and 2000 images. We have considered an equal number of images from each dataset to produce five specified subsets. Then we applied various segmentation techniques to the subsets of data. The concluded result has been described using different metrics such as precision, recall, accuracy, and F-measure, as shown in Table 1 and Figure 15.

From Table 1 and Figure 15, we observed that the simulation results of proposed frameworks, and hybridization of the *K*-means with CNN is superior to other modules in terms of the quantities parameters. Improvements in quantities parameters are clearly visible in Figure 15 and average accuracy is 85.72%, 86.06%, 87.54%, 88.64%, 92.21%, and 96.45% for FCM, *K*-means, FCM with PSO, *K*-means with PSO, FCM with CNN, and *K*-means

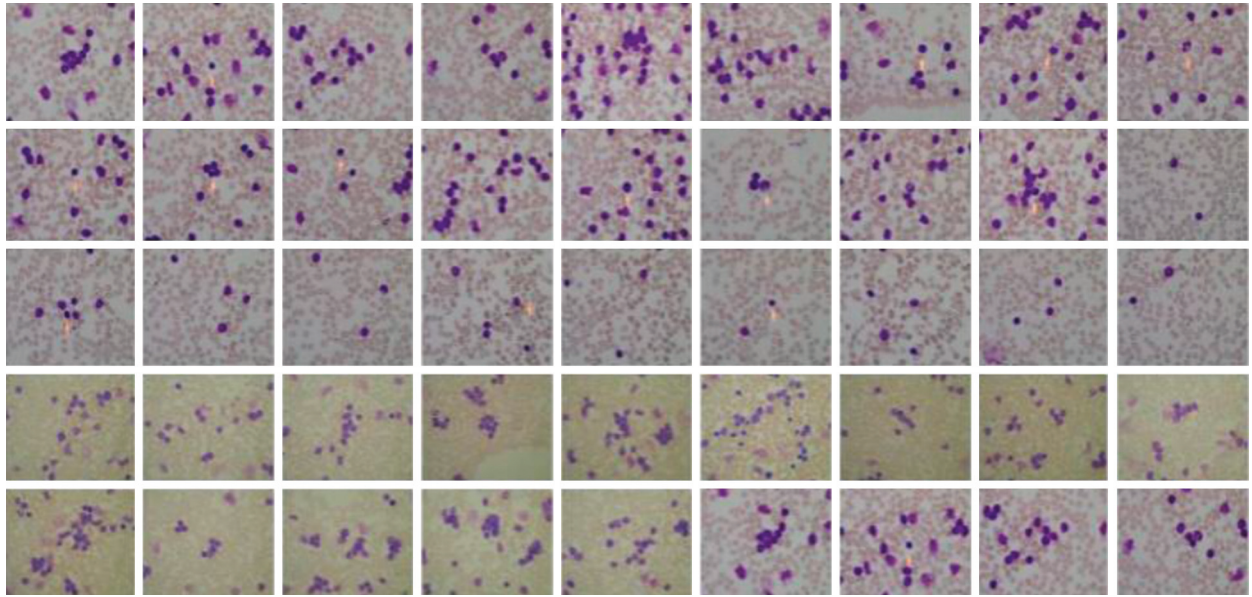


FIGURE 12: ALL-IDB dataset.

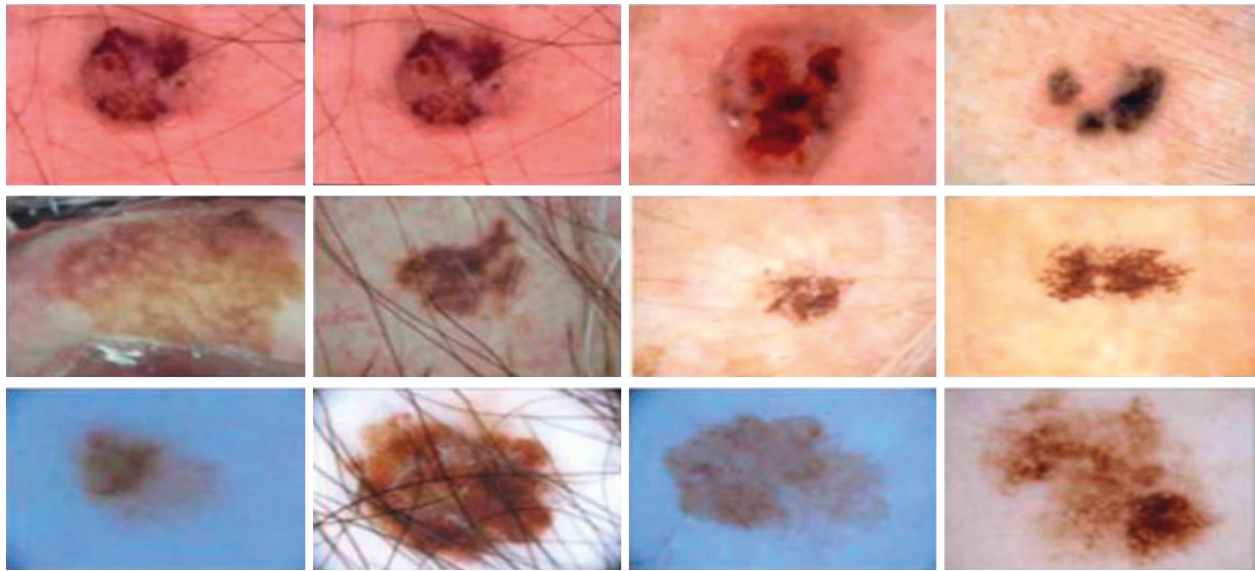


FIGURE 13: ISIC dataset.

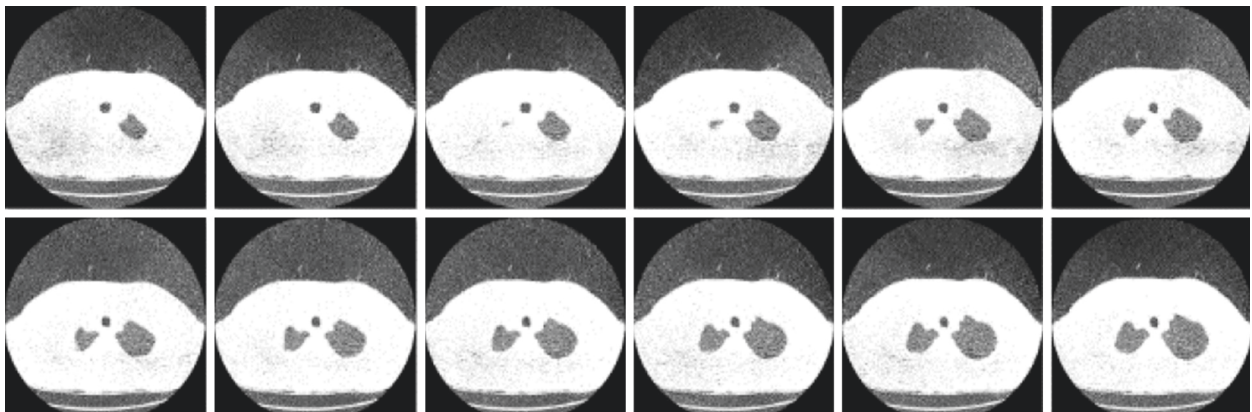
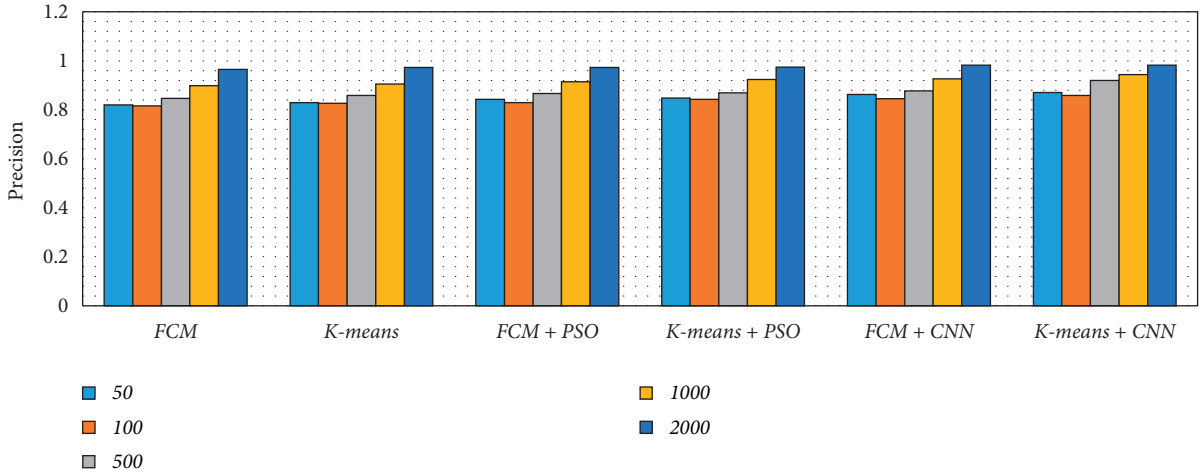


FIGURE 14: Sample of dataset CT-scan images.

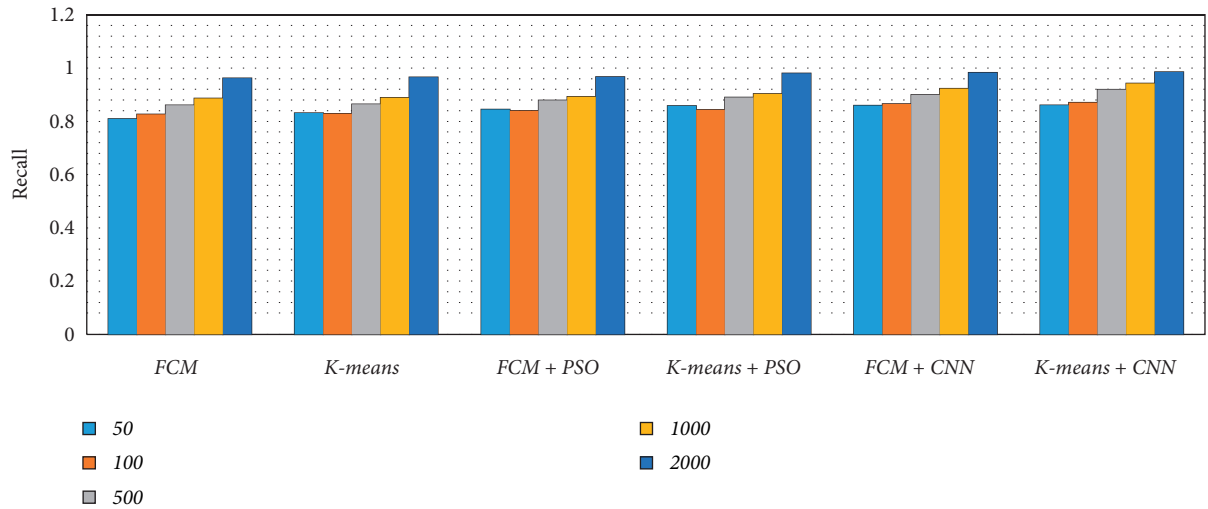
TABLE 1: Comparison of simulation results based on quantities parameters.

Parameters		FCM	<i>K</i> -means	FCM + PSO	<i>K</i> -means + PSO	FCM + CNN	<i>K</i> -means + CNN	
Precision	No. of images	50	0.8198	0.8291	0.8422	0.8474	0.8627	0.8699
		100	0.8154	0.8261	0.8295	0.8422	0.8448	0.8591
		500	0.8469	0.8583	0.8667	0.8694	0.8765	0.9197
		1000	0.8988	0.9052	0.9138	0.9237	0.9265	0.9431
		2000	0.9647	0.9725	0.9729	0.9737	0.9821	0.9821
Recall	No. of images	50	0.8097	0.8307	0.8441	0.8575	0.8599	0.8611
		100	0.8268	0.8282	0.8392	0.8430	0.8651	0.8701
		500	0.8608	0.8642	0.8792	0.8895	0.8999	0.9190
		1000	0.8864	0.8892	0.8924	0.9037	0.9232	0.9421
		2000	0.9617	0.9661	0.9673	0.9803	0.9831	0.9851
F-measure	No. of images	50	0.8147	0.8298	0.8364	0.8524	0.8612	0.8654
		100	0.8211	0.8271	0.8325	0.8426	0.8548	0.8645
		500	0.8537	0.8612	0.8686	0.8793	0.8880	0.9193
		1000	0.8925	0.8971	0.8987	0.9135	0.9248	0.9426
		2000	0.9631	0.9692	0.9698	0.9769	0.9825	0.9835
Accuracy (%)	No. of images	50	78.8307	79.5256	80.3841	81.5816	87.8346	92.8432
		100	80.3989	81.4237	82.5393	84.3495	89.8007	95.1030
		500	85.0235	86.7798	89.0326	90.0117	90.1153	96.3699
		1000	88.3359	89.2489	89.6378	89.6862	94.7604	98.2569
		2000	96.0569	96.0690	96.1101	97.6035	98.5501	99.6902
Error (%)	No. of images	50	21.1693	20.4744	13.4393	19.616	12.1654	7.1568
		100	19.6011	18.5763	10.3607	17.4607	10.1993	4.8970
		500	14.9765	13.2202	6.2858	10.9674	9.8847	3.6301
		1000	11.6641	10.7511	4.9269	10.3622	5.2396	1.7431
		2000	3.9431	3.931	2.4456	3.8899	1.4499	0.3098

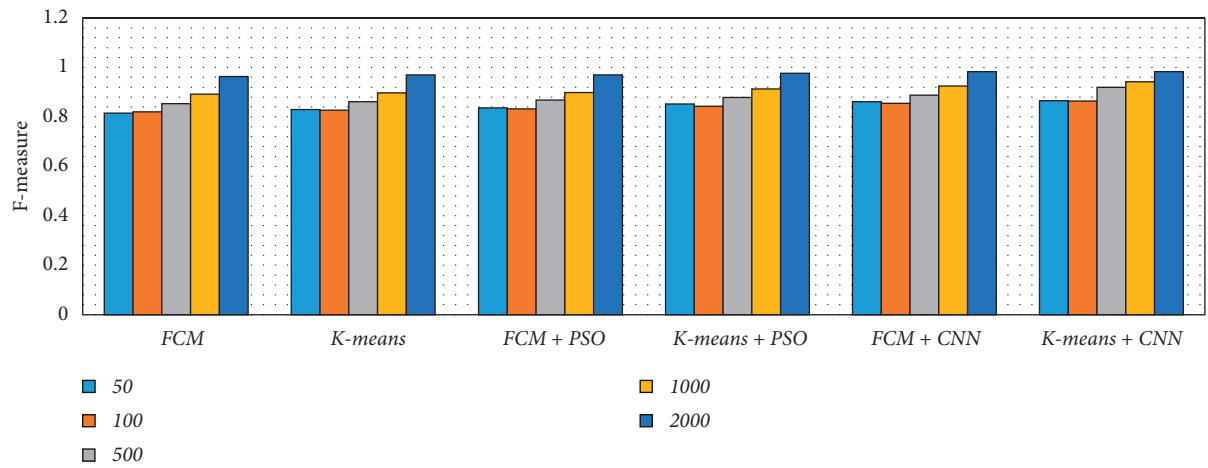


(a)

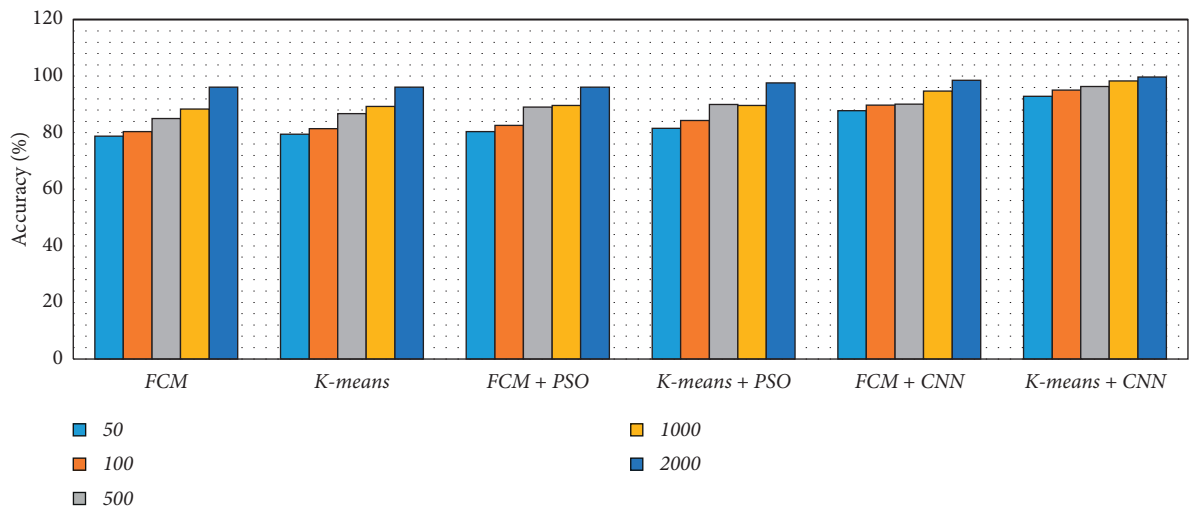
FIGURE 15: Continued.



(b)

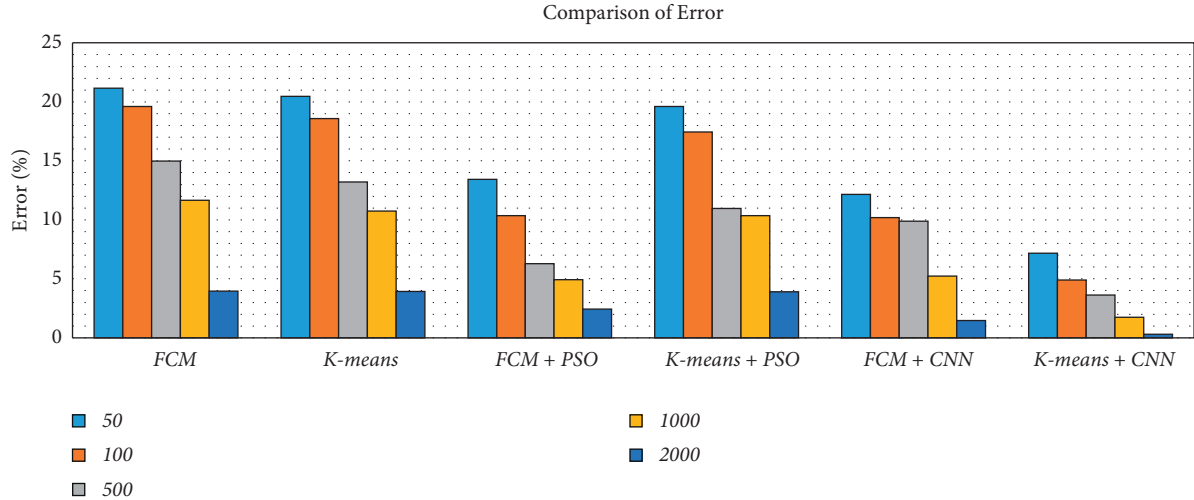


(c)



(d)

FIGURE 15: Continued.



(e)

FIGURE 15: Comparison of simulation results based on quantities parameters. (a) Precision. (b) Recall. (c) F -measure. (d) Accuracy. (e) Error.

TABLE 2: Comparison of simulation results based on similar parameters.

Parameters		FCM	K-means	FCM + PSO	K-means + PSO	FCM + CNN	K-means + CNN	
MCC	No. of images	50	0.7749	0.7917	0.8014	0.8447	0.8893	0.9706
		100	0.7792	0.7999	0.8108	0.8651	0.9157	0.9713
		500	0.8008	0.8216	0.8125	0.8858	0.9394	0.984
		1000	0.8022	0.8305	0.8587	0.8882	0.945	0.9857
		2000	0.8151	0.8656	0.8734	0.9023	0.9485	0.9979
JC	No. of images	50	0.7626	0.8085	0.8047	0.8418	0.8982	0.9668
		100	0.7853	0.8135	0.8092	0.8602	0.8993	0.9717
		500	0.7804	0.8324	0.8416	0.8834	0.9165	0.9793
		1000	0.8184	0.8406	0.8562	0.8912	0.9363	0.9867
		2000	0.8225	0.8562	0.8758	0.8957	0.9489	0.9871
CD	No. of images	50	0.7806	0.8349	0.8433	0.8897	0.8941	0.9608
		100	0.8240	0.8407	0.8528	0.8975	0.8986	0.9610
		500	0.8245	0.8431	0.8638	0.9158	0.9076	0.9695
		1000	0.8546	0.8462	0.8777	0.9221	0.9215	0.9865
		2000	0.8692	0.8879	0.8781	0.9473	0.9272	0.9888
Time	No. of images	50	7.8986	8.17185	8.21999	8.25839	8.45225	8.5419
		100	8.1251	8.12981	8.13375	8.37362	8.50933	8.53367
		500	8.41027	8.47856	8.66652	8.79246	8.83669	9.18777
		1000	8.82637	8.90073	8.92192	9.01242	9.28575	9.4042
		2000	9.60145	9.63701	9.69636	9.74011	9.76711	9.79502

with CNN, respectively. So, we can say that the effect of CNN on K -means for medical image segmentation is far better than other combinations. However, we need to validate the model based on similar parameters such as MCC, DC, JC, and computational time.

Therefore, the simulation results based on the similar values have been given in Table 2. In terms of similarities parameters, also K -means with CNN is superior to other modules for all the similarities metrics such as MCC, JC, and CD. The required time for K -means with CNN is slightly higher than other models. However, the change is extremely minute and may safely be ignored as a result. FCM with CNN is the second most successful segmentation technique based on both

quantitative and similarity metrics. Also, it is transparent that the CNN-based optimized segmentation techniques performed better than both swarm intelligence and traditional methods.

Achieving maximum accuracy is the goal of the proposed framework with a fast response to segmentation with pre-processing. The proposed framework offers an extremely self-configurable and standalone mechanism with lots of deep learning interfaces. In addition, our suggested framework generated robust models for segmenting the region of interest in various issues. So, it is a universal framework for medical image segmentation. However, to validate the efficiency of the system, we need to compare it with state-of-the-artwork based on their accuracy in Table 3.

TABLE 3: Comparison with existing works.

Previous works	Accuracy (%)
ANN (artificial neural network) for brain tumor images [17]	94.07
K-means clustering with SVM classifier for brain tumor [18]	93.00
CNN for skin lesions [23]	93.80
Swarm-based PSO inertia-weighted PSO for lung cancer [30]	95.81
SVM classifier for skin lesions [42]	94.00
FFBPNN (feed forward back propagation neural networks) for lung cancer [43]	92.60
VGG-SegNet for lung nodules [44]	99.68
K-means for brain tumor [45]	94.06
Deep learning with auxiliary task for skin lesions [46]	94.32
U-net and attention U-net for skin lesions [47]	0.913 and 0.913
Semantic segmentation for leukaemia [48]	99.10
Proposed models	Average accuracy (%)
FCM	85.72
K-means	86.61
FCM with PSO	87.54
K-means with PSO	88.64
FCM with CNN	92.21
K-means with CNN	96.45

It is clear that the previous study was conducted on a single issue such as skin lesions, brain tumors, lung cancer, or leukemia (as shown in Table 3). However, our study aims to build a universal model for dealing with all the issues. In our case, we have found an average accuracy of 96.45 with *K*-means with CNN.

5. Conclusion

In this study, we have introduced a comparative framework for medical image segmentation with traditional, swarm intelligence and convolutional neural networks as a deep learning mechanism. This framework helps to design a real-time universal medical diagnosis system for various types of images such as MRI for brain data analysis, dermoscopic for skin, microscopic for blood leukemia, and CT-scan images for lungs. Here, we present a comparative study using six different scenarios such as FCM, *K*-means, improvisation of FCM using PSO, improvisation of *K*-means with PSO, improvisation of FCM with CNN, and improvisation of *K*-means with CNN. We proved the functionality of *K*-means with CNN is a powerful hybrid mechanism that achieves 96.45% accuracy whereas, other mechanisms achieve 85.72%, 86.61%, 87.54%, 88.64%, and 92.21% FCM, *K*-means, and improvisation of FCM using PSO, improvisation of *K*-means with PSO, improvisation of FCM with CNN and improvisation of *K*-means with CNN respectively. We also found that CNN-based optimized algorithms performed well compared to optimized swarm intelligence and/or traditional methods. We expect that in the future, it will aid in the transition of medical image segmentation from research laboratories to operational or real-time applications.

Data Availability

The data are available at Brain Tumor Segmentation (BraTS) Dataset: <https://www.smir.ch/BraTS/Start2015>; Acute

Lymphoblastic Leukemia Image Database (ALL-IDB) Dataset: <https://homes.di.unimi.it/scotti/all/ISIC-2018>; Dataset: <https://challenge.isic-archive.com/data/CT-Scan>; and Dataset-<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2176079/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Y. Bengio, A. C. Courville, and P. Vincent, "Unsupervised feature learning and deep learning: a review and new perspectives," *CoRR*, vol. 5538, no. 1, 2012.
- [2] A. Tiwari, S. Srivastava, and M. Pant, "Brain tumor segmentation and classification from magnetic resonance images: review of selected methods from 2014 to 2019," *Pattern Recognition Letters*, vol. 131, pp. 244–260, 2020.
- [3] C. Bhatt, I. Kumar, V. Vijayakumar, K. U. Singh, and A. Kumar, "The state of the art of deep learning models in medical science and their challenges," *Multimedia Systems*, vol. 27, no. 4, pp. 599–613, 2021.
- [4] O. Cetin, V. Seymen, and U. Sakoglu, "Multiple sclerosis lesion detection in multimodal MRI using simple clustering-based segmentation and classification," *Informatics in Medicine Unlocked*, vol. 20, Article ID 100409, 2020.
- [5] E. Bonabeau, G. Theraulaz, and M. Dorigo, *Swarm Intelligence*, pp. 32–77, Oxford, 1999.
- [6] A. Banks, J. Vincent, and C. Anyakoha, "A review of particle swarm optimization. Part I: background and development," *Natural Computing*, vol. 6, no. 4, pp. 467–484, 2007.
- [7] P. Siva Raja and A. V. Rani, "Brain tumor classification using a hybrid deep autoencoder with Bayesian fuzzy clustering-based segmentation approach," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 1, pp. 440–453, 2020.
- [8] S. S. Chouhan, A. Kaul, and U. P. Singh, "Image segmentation using fuzzy competitive learning based counter propagation network," *Multimedia Tools and Applications*, vol. 78, no. 24, pp. 35263–35287, 2019.
- [9] J. Kennedy and R. Eberhart, "Particle swarm optimization," vol. 4, pp. 1942–1948, in *Proceedings of the*

- ICNN'95-International Conference on Neural Networks, vol. 4, pp. 1942–1948, IEEE, Perth, WA, Australia, December 1995.
- [10] D. Karaboga, “An idea based on honey bee swarm for numerical optimization,” vol. 200 Technical report-tr06, pp. 1–10, Erciyes university, engineering faculty, computer engineering department, Kayseri, Turkey, 2005.
 - [11] X. S. Yang, “Firefly algorithms for multimodal optimization,” in *International Symposium on Stochastic Algorithms*, pp. 169–178, Springer, Berlin, Heidelberg, 2009, October.
 - [12] S. S. Chouhan, A. Kaul, and U. P. Singh, “Image segmentation using computational intelligence techniques: review,” *Archives of Computational Methods in Engineering*, vol. 26, no. 3, pp. 533–596, 2019.
 - [13] S. S. Chouhan, A. Kaul, and U. P. Singh, “Soft computing approaches for image segmentation: a survey,” *Multimedia Tools and Applications*, vol. 77, no. 21, pp. 28483–28537, 2018.
 - [14] S. S. Chouhan, U. P. Singh, and S. Jain, “Automated plant leaf disease detection and classification using fuzzy based function network,” *Wireless Personal Communications*, vol. 121, no. 3, pp. 1757–1779, 2021.
 - [15] S. S. Chouhan, U. P. Singh, and S. Jain, “Web facilitated anthracnose disease segmentation from the leaf of mango tree using radial basis function (RBF) neural network,” *Wireless Personal Communications*, vol. 113, no. 2, pp. 1279–1296, 2020.
 - [16] L. Hua, Y. Gu, X. Gu, J. Xue, and T. Ni, “A novel brain MRI image segmentation method using an improved multi-view fuzzy c-means clustering algorithm,” *Frontiers in Neuroscience*, vol. 15, Article ID 662674, 2021.
 - [17] A. G. Karegowda, D. Poornima, N. Sindhu, and P. T. Bharathi, “Performance assessment of K-means, FCM, ARKFCM and PSO segmentation algorithms for MR brain tumour images,” *International Journal of Data Mining and Emerging Technologies*, vol. 8, no. 1, pp. 18–26, 2018.
 - [18] N. Arunkumar, M. A. Mohammed, M. K. Abd Ghani et al., “K-means clustering and neural network for object detecting and identifying abnormality of brain tumor,” *Soft Computing*, vol. 23, no. 19, pp. 9083–9096, 2019.
 - [19] P. S. Chander, J. Soundarya, and R. Priyadharsini, “Brain tumour detection and classification using K-means clustering and SVM classifier,” in *RITA 2018*, pp. 49–63, Springer, Berlin, Heidelberg, 2020.
 - [20] W. Zhang, Y. Wu, B. Yang, S. Hu, L. Wu, and S. Dhelim, “Overview of multi-modal brain tumor mr image segmentation,” in *Healthcare*, vol. 9, no. 8, p. 1051, MDPI, 2021, August.
 - [21] M. K. Islam, C. Kaushal, M. A. Amin et al., “A secure framework toward IoMT-assisted data collection, modeling, and classification for intelligent dermatology healthcare services,” *Contrast Media and Molecular Imaging*, vol. 2022, pp. 1–18, Article ID 6805460, 2022.
 - [22] Y. Yuan, M. Chao, and Y. C. Lo, “Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance,” *IEEE Transactions on Medical Imaging*, vol. 36, no. 9, pp. 1876–1886, 2017.
 - [23] F. Xie, J. Yang, J. Liu, Z. Jiang, Y. Zheng, and Y. Wang, “Skin lesion segmentation using high-resolution convolutional neural network,” *Computer Methods and Programs in Biomedicine*, vol. 186, Article ID 105241, 2020.
 - [24] H. Wu, S. Chen, G. Chen, W. Wang, B. Lei, and Z. Wen, “FAT-Net: feature adaptive transformers for automated skin lesion segmentation,” *Medical Image Analysis*, vol. 76, Article ID 102327, 2022.
 - [25] D. Dai, C. Dong, S. Xu et al., “Ms RED: a novel multi-scale residual encoding and decoding network for skin lesion segmentation,” *Medical Image Analysis*, vol. 75, Article ID 102293, 2022.
 - [26] P. Thapar, M. Rakhra, G. Cazzato, and M. S. Hossain, “A novel hybrid deep learning approach for skin lesion segmentation and classification,” *Journal of Healthcare Engineering*, pp. 1–21, 2022.
 - [27] N. H. M. Daud, R. A. A. Raof, M. K. Osman, and N. H. Harun, “Segmentation technique for nucleus detection in blood images for chronic leukaemia journal of physics: conference series,” *Journal of Physics: Conference Series*, vol. 1755, no. 1, Article ID 012053, 2021, February.
 - [28] P. Aiswariya and S. Manimekalai, “Global and local entropy based segmentation model for detecting leukemia in blood images,” *Annals of the Romanian Society for Cell Biology*, pp. 1914–1926, 2021.
 - [29] K. G. Dhal, J. Gálvez, S. Ray, A. Das, and S. Das, “Acute lymphoblastic leukemia image segmentation driven by stochastic fractal search,” *Multimedia Tools and Applications*, vol. 79, no. 17–18, pp. 12227–12255, 2020.
 - [30] K. Senthil Kumar, K. Venkatalakshmi, and K. Karthikeyan, “Lung cancer detection using image segmentation by means of various evolutionary algorithms,” *Computational and Mathematical Methods in Medicine*, pp. 1–16, 2019.
 - [31] W. R. P. H. van de Worp, B. van der Heyden, G. Lappas et al., “Deep learning based automated orthotopic lung tumor segmentation in whole-body mouse CT-scans,” *Cancers*, vol. 13, no. 18, p. 4585, 2021.
 - [32] F. Farheen, M. S. Shamil, N. Ibtehaz, and M. S. Rahman, “Revisiting segmentation of lung tumors from CT images,” *Computers in Biology and Medicine*, vol. 144, Article ID 105385, 2022.
 - [33] C. Kaushal, K. Kaushal, and A. Singla, “Firefly optimization-based segmentation technique to analyse medical images of breast cancer,” *International Journal of Computer Mathematics*, vol. 98, no. 7, pp. 1293–1308, 2021.
 - [34] C. Kaushal, S. Bhat, D. Koundal, and A. Singla, “Recent trends in computer assisted diagnosis (CAD) system for breast cancer diagnosis using histopathological images,” *IRBM*, vol. 40, no. 4, pp. 211–227, 2019.
 - [35] B. H. Menze, A. Jakab, S. Bauer et al., “The multimodal brain tumor image segmentation benchmark (BRATS),” *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.
 - [36] F. Scotti, “Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images,” in *Proceedings of the CIMS.A. 2005 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, pp. 96–101, Messian, Italy, July 2005.
 - [37] N. Codella, V. Rotemberg, P. Tschandl et al., “Skin Lesion Analysis toward Melanoma Detection 2018: A challenge Hosted by the International Skin Imaging Collaboration (Isic),” 2019, <https://arxiv.org/abs/1902.03368>.
 - [38] Lung Image Database, “ELCAP Public Lung Image Database,” 2021, <http://www.via.cornell.edu/databases/lungdb.html>.
 - [39] D. Chicco and G. Jurman, “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” *BMC Genomics*, vol. 21, no. 1, pp. 6–13, 2020.
 - [40] S. Jha, L. H. Son, R. Kumar, I. Priyadarshini, F. Smarandache, and H. V. Long, “Neutrosophic image segmentation with dice coefficients,” *Measurement*, vol. 134, pp. 762–772, 2019.

- [41] A. W. Setiawan, "Image segmentation metrics in skin lesion: accuracy, sensitivity, specificity, dice coefficient, Jaccard index, and Matthews correlation coefficient," in *Proceedings of the 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, pp. 97–102, IEEE, Surabaya, Indonesia, November 2020.
- [42] F. Riaz, S. Naeem, R. Nawaz, and M. Coimbra, "Active contours-based segmentation and lesion periphery analysis for characterization of skin lesions in dermoscopy images," *IEEE journal of biomedical and health informatics*, vol. 23, no. 2, pp. 489–500, 2019.
- [43] R. Arulmurugan and H. Anandakumar, "Early detection of lung cancer using wavelet feature descriptor and feed forward back propagation neural networks classifier," in *Computational Vision and Bio Inspired Computing*, pp. 103–110, Springer, Berlin, Heidelberg, 2018.
- [44] M. A. Khan, V. Rajinikanth, S. C. Satapathy et al., "VGG19 network assisted joint segmentation and classification of lung nodules in CT images," *Diagnostics*, vol. 11, no. 12, p. 2208, 2021.
- [45] A. R. Khan, S. Khan, M. Harouni, R. Abbasi, S. Iqbal, and Z. Mehmood, "Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification," *Microscopy Research and Technique*, vol. 84, no. 7, pp. 1389–1399, 2021.
- [46] L. Liu, Y. Y. Tsui, and M. Mandal, "Skin lesion segmentation using deep learning with auxiliary task," *Journal of Imaging*, vol. 7, no. 4, p. 67, 2021.
- [47] X. Tong, J. Wei, B. Sun, S. Su, Z. Zuo, and P. Wu, "ASCU-Net: attention gate, spatial and channel attention u-net for skin lesion segmentation," *Diagnostics*, vol. 11, no. 3, p. 501, 2021.
- [48] S. Saleem, J. Amin, M. Sharif, M. A. Anjum, M. Iqbal, and S. H. Wang, "A deep network designed for segmentation and classification of leukemia using fusion of the transfer learning models," *Complex & Intelligent Systems*, vol. 8, no. 4, pp. 3105–3120, 2021.

Research Article

Intangible Cultural Heritage Reproduction and Revitalization: Value Feedback, Practice, and Exploration Based on the IPA Model

Lingling Xiao ^{1,2}

¹College of History and Society, Chongqing Normal University, Chongqing 401331, Chongqing, China

²School of Marxism, Sichuan International Studies University, Chongqing 400031, Chongqing, China

Correspondence should be addressed to Lingling Xiao; 99002036@sisu.edu.cn

Received 6 July 2022; Accepted 22 July 2022; Published 9 August 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Lingling Xiao. 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.

Intangible cultural heritage has a unique value. It is very important to evaluate and discover the value of intangible cultural heritage. Therefore, referring to the relevant references of countries around the world, this paper compares the research status of countries around the world with the development trend of China's intangible cultural heritage, indicating that the network resources of intangible cultural heritage can be converted into cultural capital according to the basic theory of cultural capital and then get an economic value. The use value of intangible cultural heritage is analyzed and considered according to IPA entity model analysis, kernel density estimation (KDE), and gray correlation calculation (calculated by using IPA analysis conclusion). Among them, because intangible cultural heritage cannot be measured and verified immediately, indirect measurement verification is carried out by selecting indirect indicators such as the total number of intangible cultural heritage and the development trend of intangible cultural heritage tourism. Based on the gray correlation analysis of the total number and relative density of national intangible cultural heritage and the intangible cultural heritage tourism and related industrial chains, the following proposals for the development trend of intangible cultural heritage are clearly put forward as follows: (1) Further excavate the meaning of "non-material property tourism" and increase the scope of development and design of cultural and art tourism network resources. (2) Make full use of the guiding effect of government departments on "intangible cultural heritage tourism," and improve the normalization of the combination of culture, art, and tourism. (3) Expand the whole industrial chain of "intangible cultural heritage tourism" in an orderly manner, and promote the concept of coordinated and sustainable development of cultural and art tourism.

1. Introduction

People's intangible cultural heritage (hereinafter referred to as intangible cultural heritage) is a unique way of key historical and cultural heritage in the world. It has a particularly key and unique meaning of world history, culture, art, plastic arts, and cultural aesthetics. It is also a precious cultural and artistic capital of global human society. It is the continuation of the spiritual essence of "being" in the history of global human development and the deep imprint of the nation. It is the representative of the national flag and national identity. It is also a symbol of national flag and national identity. Unlike intangible cultural heritage, its detailed existence depends on the initial media created by people and is passed

down from generation to generation according to oral tradition. Therefore, scientific research on intangible cultural heritage has key practical significance for both the country and the nation [1].

The development and design of intangible cultural heritage is in line with the provisions of the period, the provisions of independent innovation in the transformation and development of intangible cultural heritage and tourism, and the feasibility analysis of exploring "poetry" and "distance" in the collaborative development trend of culture and tourism [2]. Various key intangible spiritual essence historical and cultural heritages are the specific content of rich and excellent corporate culture accumulated and inherited by the Chinese nation in the

long-term material production and social life and are the key components of the 5000 year Chinese civilization. Systematic scientific research is carried out according to the style of taking the vast and thick intangible cultural heritage as the theme, integrating a variety of folk culture “intangible property” into the development of tourism resources, and taking this as a breakthrough, collaboration is carried out in the fields of current policies, network resources, sales market, and industrial chain innovation to assist in the maintenance and utilization of intangible cultural heritage, so that we can not only spread, inherit, and sustain Chinese traditional culture, At the same time, it also creates new industrial chains, new development trends, and new driving forces for tourism, and promotes the further coexistence of intangible cultural heritage and tourism. In the new era of “cultural tourism integration,” the cultural tourism industry chain is developing from a new perspective. In the new era of “integration of culture and tourism,” taking “intangible cultural heritage” and “tourism industry” as relevant research objects, this paper deeply ponders the coordination between the inheritance of “intangible cultural heritage” and the independent innovation of holiday tourism. In the new era of “cultural tourism integration,” we take “intangible cultural heritage” and “the tourism industry” as relevant research objects. Considering the inheritance of “intangible cultural heritage” and the shared development of independent innovation of vacation tourism, we are concerned about the traditional cultural extension of “intangible cultural heritage,” the collaborative requirements of independent innovation of vacation tourism, and everyone’s spiritual needs. The optimization, allocation, and reconstruction of cultural tourism network resources are completed according to the complementary methods of continuously infiltrating and expanding the maintenance and utilization of intangible cultural heritage and the reform and innovation of the tourism industry. In accordance with the complementary methods of continuously infiltrating and expanding the maintenance and utilization of intangible cultural heritage and the reform and innovation of the tourism industry, we will improve the allocation and reconstruction of cultural tourism network resources, and basically build a new comprehensive management system with the dual role of promoting intangible cultural heritage and the development trend of tourism development. It is an urgent daily task to explore the independent innovation and integration of “intangible cultural heritage vacation tourism,” reproduce cultural tourism, innovate independently in a modern environment, show the traditional culture to the public, and explore effective ways for the coupling and shared development of intangible cultural heritage and tourism development. This is also an urgent daily task [3].

Therefore, this paper discusses the value of intangible cultural heritage through certain research methods to assess the value and index measurement of intangible cultural

heritage, which is measured and reflected through the economic value of the tourism industry.

2. Research Background

Foreign scholars and experts have explored the two-way relationship between intangible cultural heritage and the tourism industry through different perspectives of positive and negative aspects. On the positive side, researchers analyze and explore through the correlation between the two and their positive influence on each other. In terms of intangible cultural heritage and tourism industry correlation, Nasser explored urban cultural heritage preservation and sustainable tourism development from the perspective of effective correlation and mutual promotion between intangible cultural heritage and the tourism industry [4]. Loulanski et al. argued that intangible cultural heritage and the tourism industry are mutually reinforcing relationships and that the orderly integration of intangible cultural heritage and the tourism industry have sustainable viability [5]. Timothy argued that the mutually supportive development of the cultural heritage sector and the tourism industry are an emerging trend in the integration of contemporary cultural economy [6]. Throsby explores the path of sustainable economic and cultural development by constructing the “golden rule” for the codirectional development of intangible cultural heritage and the tourism industry with relationship building as the main research point [7]. Nega pointed out that management and value evaluation in cultural tourism and intangible cultural heritage are correlated, and they not only support each other but also enhance each other’s value, and they analyzed and studied the case of the Lalibela Ashenda Festival in Ethiopia [8].

In domestic theoretical research, scholars and experts have evaluated the relationship between intangible cultural heritage preservation and the tourism industry in both positive and negative directions. In terms of positive effects, Lin argued that in the current context of experience economy, tourism experience of intangible cultural heritage as an interactive model is an effective path for intangible cultural heritage preservation and leisure tourism development [9]. Wang proposed the concept of “authentic tourism” and believed that the positive interaction between intangible cultural heritage and tourism could help to maximize the satisfaction of the demand and experience value of public tourism [10]. According to Wang, intangible cultural heritage and tourism have a natural relationship, and both intangible cultural heritage and tourism activities are essentially cultural in nature, with tourists pursuing cultural goals and intangible cultural heritage itself being the carrier of culture. The intangible cultural heritage, which is constantly evolving, has its own characteristics, and its new content is the driving force behind the continued vitality of tourism products and businesses. While enriching the content of tourism, it is also an important path to promote the protection and transmission of intangible cultural heritage [11].

3. Research Methodology and Materials

3.1. Main Theories

3.1.1. Intangible Cultural Heritage. The cultural heritage defined by China and the world can generally be divided into two categories: “tangible” and “intangible” cultural heritage. In the hearts of the vast majority of people, “tangible cultural heritage” is “the chemical species of cultural heritage,” such as precious cultural relics, pottery, and relics [12]. “Intangible cultural heritage” means that there is no entity but not limited to performance, cultural knowledge, and professional ability. According to the form of information, they are divided into “tangible” and “intangible” cultural heritage categories [13]. Tangible cultural heritage refers to movable property mortgage, real estate, and key precious cultural relics in the historical period. Intangible cultural heritage refers to some traditional main forms of material civilization theme activities and forms that exist in a designated intangible form and are in direct contact with everyone’s daily life at that time.

The definition of “intangible cultural heritage” (intangible ICH) was introduced into Japan by the last “law on the protection of cultural and artistic assets” implemented worldwide in Japan. Some cultural heritage with no physical shape is called “intangible ICH,” which is a kind of cultural heritage. In October, 2003, the concept of “intangible cultural heritage” was gradually accepted and recognized by people all over the world [14]. At the same time, in order to better understand, the regulations on safeguarding intangible cultural heritage based on UNESCO documents have been implemented. The list of national intangible oral cultural heritage is divided into the following five categories: firstly, oral short stories; secondly, English composition model forms; thirdly, art performance, social development customs, and plastic arts; Fourthly, traditional handicrafts; Fifthly, the understanding and practice of nature. However, due to the different domestic situations, the National Intangible Cultural Heritage Federation has clearly put forward many new opinions on the classification of intangible cultural heritage.

The evolution of intangible cultural heritage maintenance [15] is shown in Figure 1.

As shown in Figure 2 below, Chinese scholars have two symbolic views on the definition of “intangible cultural heritage:” one thinks that “intangible cultural heritage” includes various traditional Chinese cultural methods, such as folk-themed activities, traditional performing arts, and cultural indoor space; the other thinks that “intangible cultural heritage” includes various traditional Chinese cultural methods, such as folk-themed activities, traditional performing arts, and cultural interior space. Second, “intangible cultural heritage” is a kind of cultural heritage that is spread from generation to generation according to oral and spiritual essence. According to the concept of the national standards of the United Nations Organization for culture, education, science, rationality, and culture and art institutions, and the norms of the opinions and suggestions on strengthening the maintenance of China’s intangible

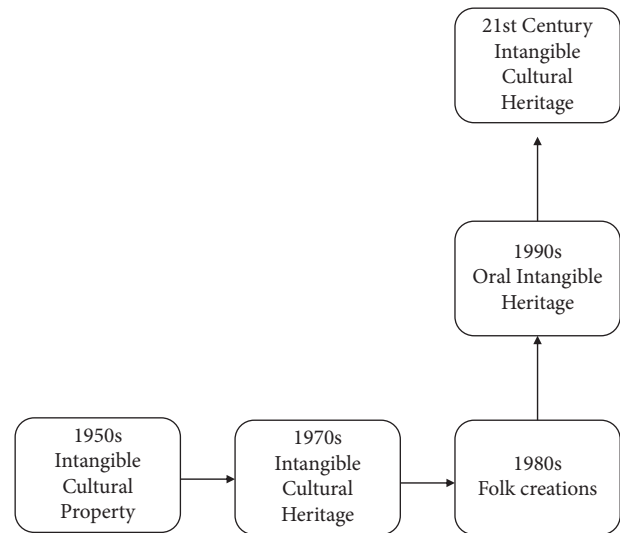


FIGURE 1: Evolution of the concept of intangible cultural heritage.

cultural heritage announced and implemented by the former Ministry of culture in 2005, this scholar clearly puts forward the main definitions of intangible cultural heritage, which are a part of the whole process of recording the cultural heritage of relevant communities, groups, institutions, and residents. Special tools, objectives, and venues for the examination of professional knowledge, professional skills, social development and practical skills, and related common sense. The implication category of Chinese national intangible oral cultural heritage works includes the following six primary property components [16].

According to the specific heritage categories contained in the document catalogue of national protected intangible cultural heritage in the current standard classification of Chinese folk intangible cultural heritage, Chinese folk plastic arts can be effectively divided into the following ten categories: folk literature, folk music, folk dance, folk art, traditional handicraft, traditional drama, opera, traditional acrobatic performance activities and athletics, traditional medicine, and folklore. After considering the current research achievements of intangible cultural heritage around the world in an all-round way, many achievements have been made closely around the regulations on safeguarding intangible cultural heritage. After continuous scientific research, with the rapid development of society and the gradual deepening of relevant discussions in countries around the world, people have further enriched and promoted their understanding of the definition and meaning of intangible cultural heritage.

3.1.2. Cultural Capital Theory. In the 1980s, the European sociologist Bourdieu, in the preface chapter of his book “Forms of Capital,” for the first time, creatively reargued and defined a new set of theoretical foundations and theoretically basic material concepts of the category of material capital, “cultural capital” [17]. He thus proposed that material capital in the broad sense that can be subdivided into three main theoretically fundamental material forms: economic capital,

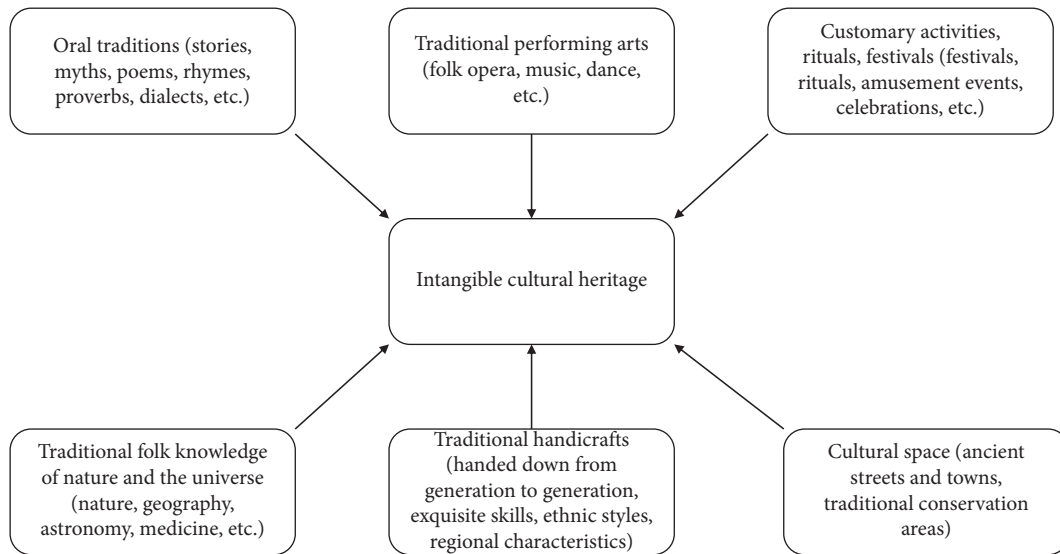


FIGURE 2: The scope covered by intangible cultural heritage in China.

cultural capital, and material social capital. Cultural capital can be further divided into the following three main material forms of culture: first, a subjective state of human existence, expressed as a state of mind and a state of spiritual insistence on the meaning of human life; second, a systemic objective state, in which the information material of cultural goods (pictures, books, dictionaries, tools, machines, etc.) with a specific systemic category of content is the objective state (tools, machines, etc.) as an objective basis of a form of information carried out and presented; third, system state information should also all be presented in its entirety in the form of dynamic information describing changes in the objective state of information. Culture itself has a value, and this intangible value can of course be transformed into a tangible economic value. Cultural capital, as objective existence, and subjective economic capital can be unified to form cultural products. The cultural capital in Bourdieu's theoretical elaboration is the theoretical basis for the productive conservation and industrial development of intangible cultural heritage. On the one hand, cultural capital guides the rational production and consumption activities of people, and on the other hand, it determines the innovation of ideas and the constant adaptation to the changing needs of people. The introduction of Bourdieu's theory of cultural capital has sparked a hot topic of discussion about culture and its production in the Western academic world. David Crosby, an Australian scholar, explains that, in addition, "cultural capital is the accumulation of cultural values in the form of wealth." In addition, he himself tries to divide "cultural capital" into "tangible cultural capital" and "intangible cultural capital" [18]. The process of accumulation of tangible resources and tangible cultural capital generally occurs in buildings, sites, art museums, and a large number of other cultural artifacts, such as paintings, sculptures, and many other cultural artifacts that are widely available in the form of their private objects and will eventually be widely attributed with a specific cultural meaning (often referred to as "cultural heritage"). Intangible cultural capital consists of

a set of ideas, practices, beliefs, traditions, and values that are consistent with a particular group of people. The important historical and cultural values and their connotations contained in the composition of intangible cultural capital are often in direct correspondence with each other or with other historical and economic values of human beings. For example, the existing rich ancient cultural, musical and artistic texts, modern historical and literary works, the folk wealth and accumulation, cultural customs (habits) and various spiritual beliefs, or various cultural languages, have an extremely direct and extensive deep and lasting political and spiritual cultural value, but often have no direct political and economic value at all, because at times they can no longer be truly valued as a material wealth value to be traded for the social value. However, the various commercial and service activities and product flows associated with these direct flows of material intangible wealth and tangible cultural capital can create a direct cultural value and a potential economic value for society, respectively. The "intangible culture" is an important characteristic and representation of the "intangible culture" part of the "intangible cultural heritage," as shown in Figure 3.

3.2. Research Methodology

3.2.1. IPA Analysis Method. IPA analysis (importance performance analysis) is a method to accurately measure and objectively analyze the intrinsic importance of a specific product characteristic or a specific service content beyond the subjective expectations of customers, and to observe the intrinsic differences between the levels of performance indicators and technical performance. It is a quantitative analysis method to find the possible points to improve productivity or further improve the target customer satisfaction. It was first proposed by Martilla et al. [19]. Evans et al. [20] proposed for the first time in China to formally apply the IPA model and analysis method tool techniques to

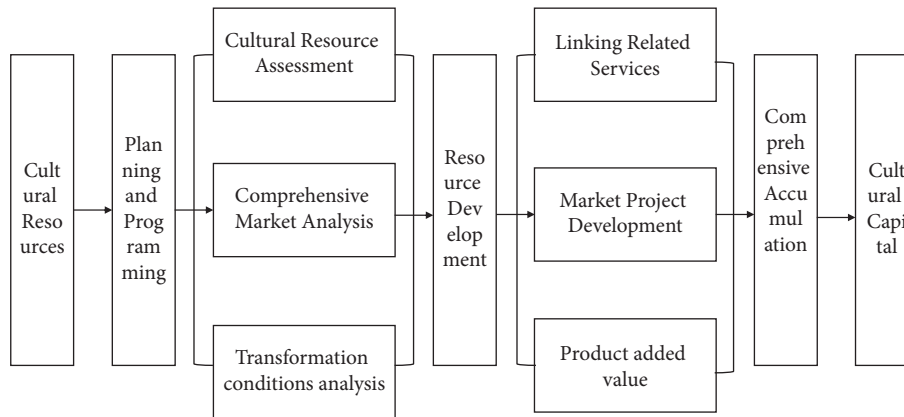


FIGURE 3: The transformation path of cultural resources into cultural capital.

develop basic research in the field of domestic cultural tourism. In recent years, IPA analysis method tool technology will mainly apply the research work to improve the basic service resources supply and its quality in China's tourism and the culture industry, the value of tourists' destination perception experience, tourism service industry destination image value enhancement and other three factors related to the basic comprehensive method research. For example, the National Natural Science Foundation of China Geng [21] project and other items. Based on the perspective of China and ethnic regions, a hypothetical evaluation model was constructed as a Chinese tourism service industry of service and provision of service quality dimension and satisfaction evaluation dimension index differences. Through the empirical questionnaire analysis questionnaire data and IPA data comprehensive analysis, to explore the most specific difference between tourism hotels and services and tourist quality satisfaction difference indicators in the three countries mentioned above different evaluation dimensions data level between the most specific difference cause analysis and difference analysis method importance. The research backbone of the group such as Zunan [22] proposed to study and build up a model system for measuring the customer perception and an experience index of tourism hotel chain B&B based on the research results of hotel B&B tourist perception and the experience evaluation value model, using separate methods such as IPA regression analysis, factor regression analysis, and equal weighting method to investigate the tourist perception between various types of economic tourism hotel chain B&B at home and abroad. The quantitative and empirical comparative evaluation of the differences in indicators of experience evaluation was carried out, and a series of tourism-related planning strategies and suggestions for improving the perception indicators of accommodation tourists and destination experience perception indicators were proposed. The project of Jing Xuli et al. research team [23] mainly used the web text analysis model and the IPA model to analyze and compare some subtle cognitive differences based on individual tourists' overall perception and way perception of cultural image characteristics of the same global tourism route destination, with a view to help China to optimize

international tourism and cultural destination image resources for international tourism resources strategy [24].

This paper focuses on the valuation and index measurement of intangible cultural heritage using IPA analysis also through the indirect measurement of the tourism value. The measured values are not directly represented in the article but in the calculated values of the gray correlation, which is calculated by IPA analysis. Therefore, this paper emphasizes IPA analysis to fully evaluate and measure the relationship between ICH and the tourism industry and analyze the value of ICH.

3.2.2. Multimodel Combined Coupling Evaluation Method.

A three-dimensional evaluation of structural coupling, dynamic coupling, and stage coupling is carried out for the intangible cultural heritage and tourism industry systems by combining multiple models. Based on the gray correlation model, the structural coupling of the two systems is analyzed by the CPA (based on IPA) model; the entropy change model is used to explore the coupling relationship between the two systems in the dynamic development direction; the coupling coordination degree model is used to evaluate the coupling development level of the two systems, and the different spatial and temporal development stages of the coupling are classified. The coupling coordination degree model is used to evaluate the coupling development level of the two systems, and to classify the different development stages of the coupling and the synchronization of the two systems. The path and method of overall coupling evaluation are provided for the empirical research.

3.3. Equation Theory

3.3.1. *Kernel Density Estimation (KDE)*. China's intangible cultural heritage is starry and huge, but due to the different conditions of different cultural elements, the density of projects in each region is unevenly distributed from the situation of regional excavation and collation declaration, and some regions are still blank areas. Taking the representative national intangible cultural heritage (data source: China's intangible cultural heritage network) as the

measurement object, we try to map the density of national intangible cultural heritage items by Kernel density estimation (KDE), a spatial analysis tool in ArcGIS software, according to the probability theory.

Kernel density is defined as: set x_1, \dots, x_n the point set is a sample f drawn from the overall of the distribution density f function as, x estimated $f(x)$ at a point. The Rosenblatt-Parzen kernel estimation is usually used, and its calculation formula is

$$f_n(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right). \quad (1)$$

In (1), the $kx - x_1/h$ formula is the kernel $h>0$ function and $x - x_1$ is the broadband, which indicates the distance from the valuation point to the event. The main factors that affect KDE are the mathematical form of the kernel function and the value of the bandwidth.

3.3.2. Gray Correlation Degree Calculation Method. The gray correlation analysis method is used to analyze the gray correlation degree of various intangible and cultural heritage in the tourism service industry and its development with related multi-industries, which can reflect more objectively and truly the gray cross-correlation and relationship between various intangible or cultural heritage and tourism-related industry development and other multi-related industry development. The data are extracted from 2000 to 2009 and from 2010 to 2019. The gray correlation model was constructed by extracting two sets of relevant data (based on the China Intangible Cultural Heritage Tourism Statistical Yearbook and relevant data from the National Bureau of Statistics) from 2000 to 2009 and from 2010 to 2019. The numerical potential of the gray correlation model is based on the IPA analysis method, and the values calculated by the IPA analysis method are used to calculate and analyze the gray correlation degree. x_i , m is the selected relevant industry, and n is the time series of the study, where $X_i = (X_i(1), X_i(2), \dots, X_i(k))$ is the data of the selected industry related to tourism.

The calculation formula is as follows:

$$x_i(k) = \frac{x_i(k)}{x_i(1)}, k = 1, 2, \dots, n, i = 1, 2, \dots, m. \quad (2)$$

After normalization, the obtained x_i matrix is

$$x_{i=}(x_i(1), x_i(2), \dots, x_i(k)) i = 0, 1, 2, \dots, m. \quad (3)$$

Calculate the difference series as

$$\Delta_i = (\Delta_i(1), \Delta_i(2), \dots, \Delta_i(k)) i = 0, 1, 2, \dots, m. \quad (4)$$

Calculate the two extreme differences as

$$M = \max_i \max_k \Delta_i(k), m = \min_i \min_k \Delta_i(k). \quad (5)$$

Calculate the correlation coefficient as

$$\gamma_{\partial}(k) = \frac{m + \zeta M}{\Delta_i(k) + \zeta M}, \zeta \in (0, 1), i = 1, 2, \dots, i = 1, 2, \dots, m. \quad (6)$$

Calculate the gray correlation coefficient as

$$\gamma_{\partial}(k) = \frac{1}{n} \sum_{k=1}^n \gamma_i(k), i = 1, 2, \dots, m. \quad (7)$$

4. Results and Discussion

4.1. Status of Intangible Cultural Heritage. Intangible cultural heritage, as a part of the excellent Chinese traditional culture, is long-lasting and new, and has great cultural and scientific values. Strengthening the cognition and development of intangible cultural heritage, is the identification of traditional cultural genes and national cultural lineage, as well as the process of awakening people's cultural self-reflection and enhancing national cultural self-confidence. The protection and development of intangible cultural heritage in China began at the beginning of this century, and although it started late, it is highly valued by the Party and the State as an important cultural work in the new era. China's intangible cultural heritage has gone through three stages of development: the initial stage (2001–2005), the development stage (2006–2010), and the prosperity stage (2010–present), and the amount of intangible cultural heritage items, the number of certified bearers, and the cultural ecology. The number of items, the number of certified bearers, and the number of protected areas are now at the forefront of the development of intangible cultural heritage in the world. China's intangible cultural heritage has been transformed from a "grassroots culture" to a national "cultural symbol" and has become a synonym for "national cultural genes." In 2006 and 2007, the Anhui province published and confirmed the list of the first batch of local representative provincial intangible cultural heritage protection and inheritance projects at the provincial level and the third batch at the county level through China's third review. In 2011 and 2014, according to the State Council through the Ministry of Engineering and Construction of the State Ministry of Environmental Protection approved the review determined and at the same time confirmed the publication and generation of the third and fourth batches of China's Anhui two representative national list of the first batch of intangible cultural heritage of the Anhui province. In 2012 and 2018, the State Administration of Cultural Heritage has jointly announced the identification in order to protect the worldwide fourth and fifth batches of two national representative projects of human significant intangible cultural heritage representative techniques and inheritance of regional representative people. Among them, China's two national major intangible cultural heritage of humanity projects and the cumulative scale of the total number of conservation achievements have been ranked on the world's first in both successive countries. At this stage, China's intangible heritage excavation projects have not only shown a prosperous

trend but also made certain achievements in productive conservation and application development.

As shown in Figure 4, as of 2020, the categories of world intangible cultural heritage declared by China have been inscribed by the United Nations on the list of Intangible Cultural Heritage of Humanity with a total of 39 items, which also include 32 items in the international general endangered list and a total of 7 items in the international endangered list for the types of world intangible cultural heritage that urgently need international protection list, thus, China is currently the country with the largest number of intangible heritage. Therefore, China is the country with the largest number of masterpieces of intangible heritage. At the same time, China has built a complete four-level protection and a certification system at the national, provincial, municipal, and county (district) levels, and government departments at all levels have been working in parallel to excavate local intangible cultural heritage types of projects with cultural characteristics and a high aesthetic value, with remarkable results. After more than a decade of accumulation, China now has 1,530 national-level intangible cultural heritages, covering 10 intangible cultural heritage categories. Meanwhile, China has 5,219 provincial-level intangible cultural heritages, 10,331 provincial-level intangible cultural heritages, and 72,776 provincial-level intangible cultural heritages at the county (district) level. The number of intangible cultural heritage at all levels is among the highest in the world.

As shown in Figure 5, on the basis of the ten categories of intangible cultural heritage, as of December 2020, the Ministry of Culture's Office of the Special Working Group of the Ministry of Culture of China (now renamed the Ministry of Culture and Tourism) was established to date has initially identified a total of approximately 3,068 people as the main inheritors and representative persons of the representative units of the first five major batches of projects at the national level or major categories of intangible cultural heritage projects. In addition, in China this year, other provinces and municipalities across the country have announced or recognized a total of about 14,928 local municipalities and counties below the level of representative skills excellent inheritors of intangible cultural heritage, together building a genealogy of Chinese cultural heritage. From 2015, the field of culture and tourism, the implementation of "China's intangible cultural heritage inheritance of population training program," the use of academic and teaching resources to support, focusing on minority areas and remote and poor areas, in the inheritance of traditional process to carry out "strong foundation," and broaden their horizons increase. The program focuses on ethnic minority areas and remote and poor areas to "strengthen the foundation," broaden horizons, and increase knowledge in the process of passing down traditions. At the same time, the inheritors are included in traditional craft workstations and workshops of "intangible cultural heritage," and the enthusiasm, initiative, and creativity of the inheritors are continuously mobilized. According to statistics, by the end of 2020, 121 institutions across the country, in which the research and training programs are held, have held a total of more than 850

training sessions of various types, and more than 33,000 participants have been trained, in addition to the content of local regions and related forms of extended training, a total of more than 100,000 people covering the group of inheritors.

As shown in Figure 6, China, based on its own national conditions, has creatively taken the representative "intangible cultural heritage" as the core, with the purpose of reshaping the cultural system and the soil of "intangible cultural heritage" as the core. At the same time, 168 provincial-level cultural ecological protection zones have been established, including Zulu culture, Anqing theater culture, Sli County woodblock printing culture, Zhangnan traditional cotton weaving technique culture, and Jiangyan Qingming custom culture. China's national and provincial cultural ecological reserves show a benign trend of year-on-year growth. China has become one of the countries with the largest number of cultural ecological reserves and the largest area in the world.

4.2. Related Conclusions

4.2.1. Schematic Diagram of the Distribution Number and Density of National Intangible Cultural Heritage. China's intangible cultural heritage is splendid and huge, but because the situation varies from one cultural element to another, the density of projects in each region is unevenly distributed from the situation of regional excavation and collation declaration, and some regions are still blank areas. Taking the representative national-level intangible cultural heritage (data source: China Intangible Cultural Heritage Network) as the measurement object, we tried to map the density of national-level intangible cultural heritage items through the Kernel density estimation (KDE) method of the spatial analysis tool in ArcGIS software according to the probability theory. The distribution number and density of national-level intangible cultural heritage are schematically shown in Figure 7.

4.2.2. Gray Correlation between the Intangible Cultural Heritage Tourism Industry and Related Industries. As shown in Figure 8, based on the IPA model analysis method, this paper obtains the gray correlation degree between each industry and the intangible cultural heritage tourism industry in two stages from 2000 to 2009 and from 2010 to 2019 by calculating the absolute difference and the time series correlation coefficient. It can be seen that the association density of the traditional type of closely related industries has been increasing, with the restaurant industry increasing from 0.6714 to 0.9815, the association of the accommodation industry increasing from 0.6404 to 0.9191, the transportation industry increasing from 0.5709 to 0.9264, and the landscape industry increasing from 0.7295 to 0.8595. At the same time, the original low degree of association with the intangible cultural heritage tourism industry, with the tourism industry associated with the expansion of the boundaries of the close degree is also gradually strengthened, such as with the continuous

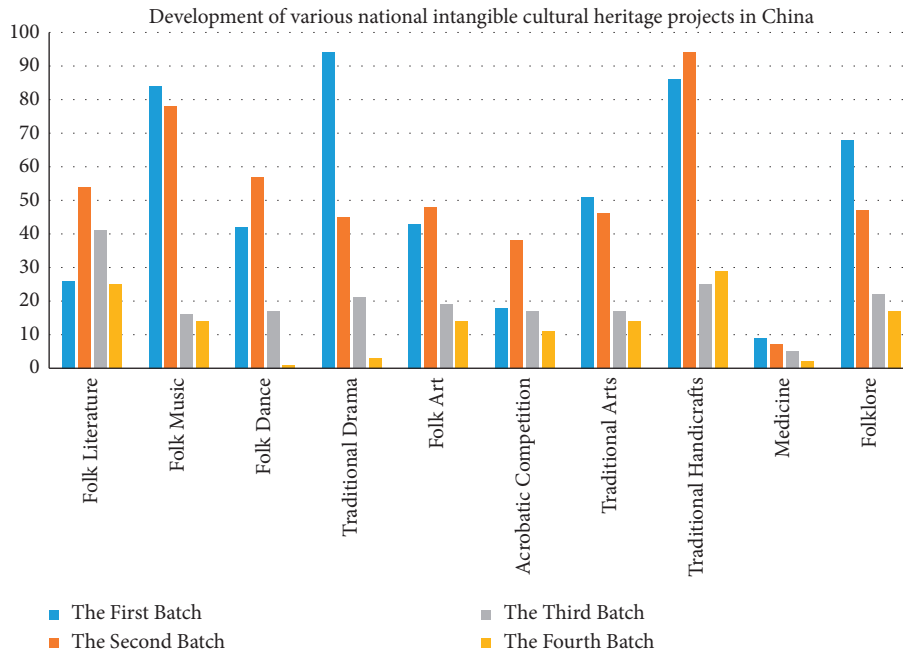


FIGURE 4: Development of various types of national intangible cultural heritage projects in China.

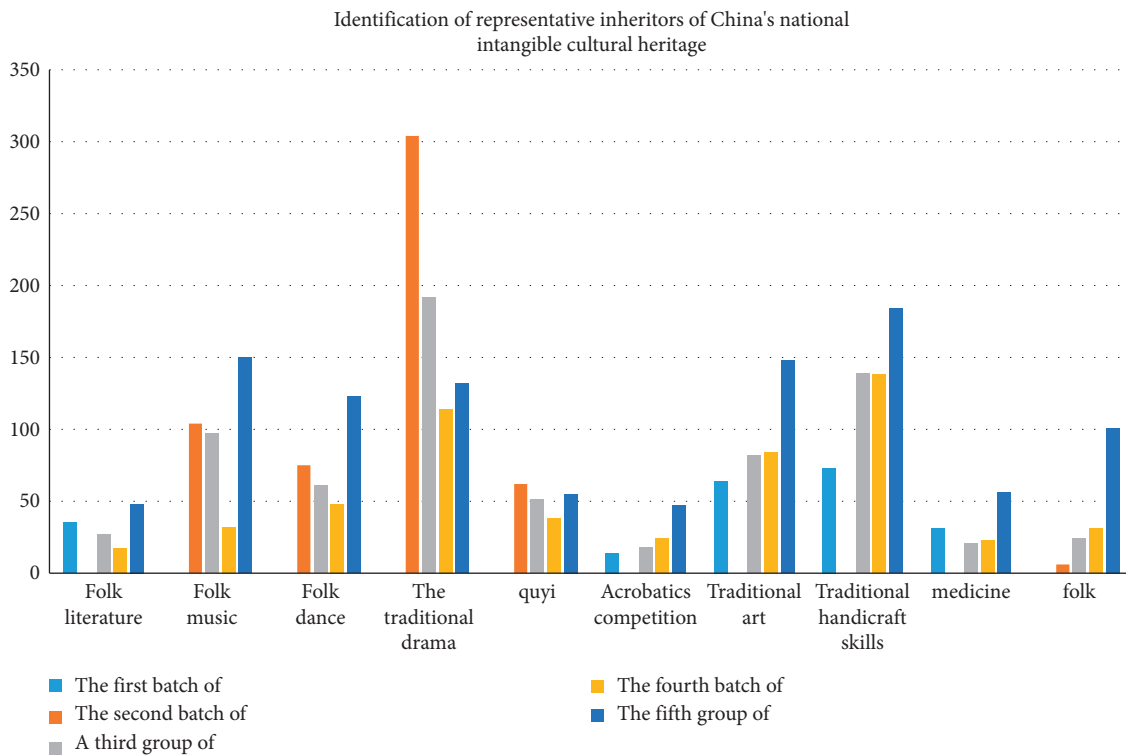


FIGURE 5: The recognition of representative inheritors of various types of intangible cultural heritage at the national level in China.

construction of tourism services and infrastructure, tourism and construction industry association from 0.7305 to 0.8399 along with the rise of new business models, such as agricultural tourism, rural tourism, and tourism and agriculture between the gray. The correlation between tourism and agriculture increased from 0.6589 to 0.9402 with the

proposal and development of the national cultural tourism integration plan, the correlation between tourism and culture and the art industry also increased from 0.7926 to 0.9427, and the effect of tourism linking multiple industries became increasingly obvious, and gradually formed a positive interaction between industries.

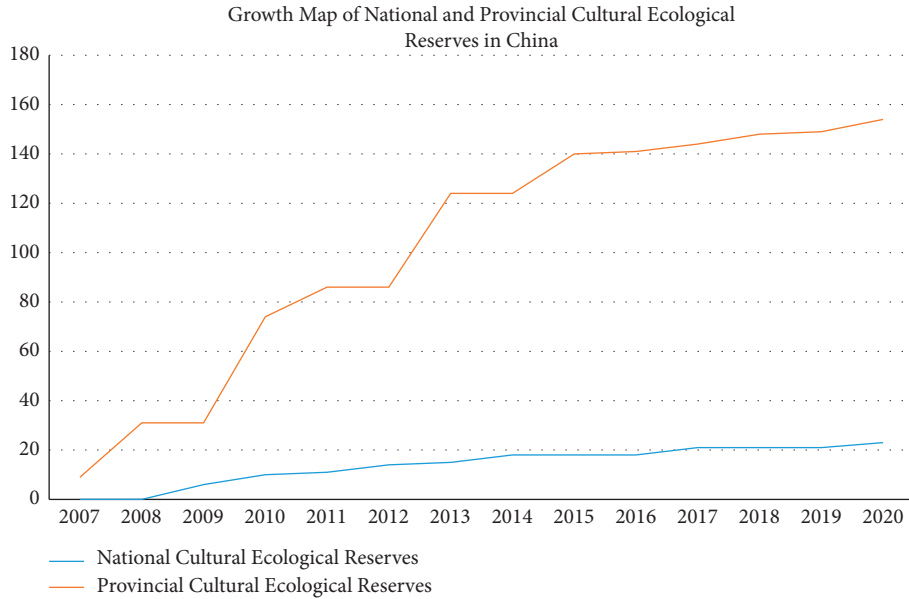


FIGURE 6: Growth of national, provincial, cultural, and ecological reserves in China.

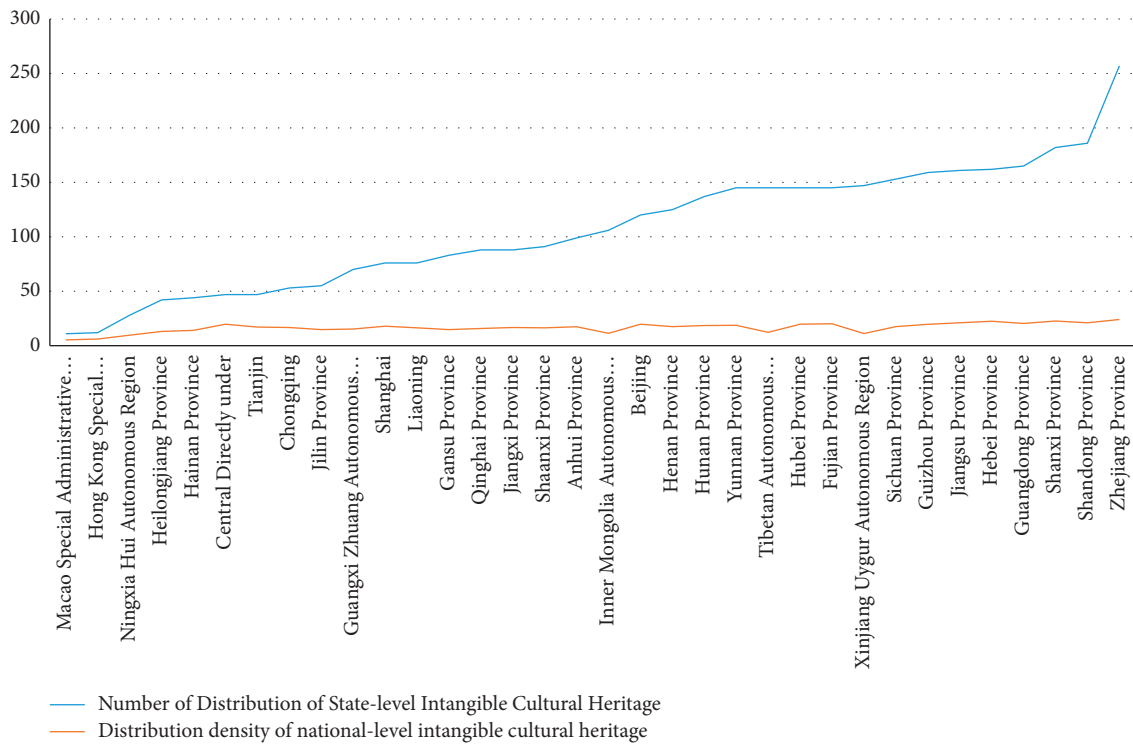


FIGURE 7: Schematic diagram of the distribution quantity and density of national-level intangible cultural heritage.

4.3. Suggestions for Coupling Development of Intangible Cultural Heritage and the Tourism Industry Are Discussed

4.3.1. Deeply Excavate the Connotation of “Intangible Cultural Heritage + Tourism” and Increase the Development of Cultural and Tourism Resources. Based on the analysis of the above empirical results, the coupled and coordinated

development should be promoted on the basis of the integration and joint development of the resources of “intangible cultural heritage+tourism,” and the image of cultural tourism should be enhanced through the deep excavation of the connotation of “intangible cultural heritage+tourism.” To further explore the connotation and related contents of intangible cultural heritage and tourism common source in each local area, and to explore the

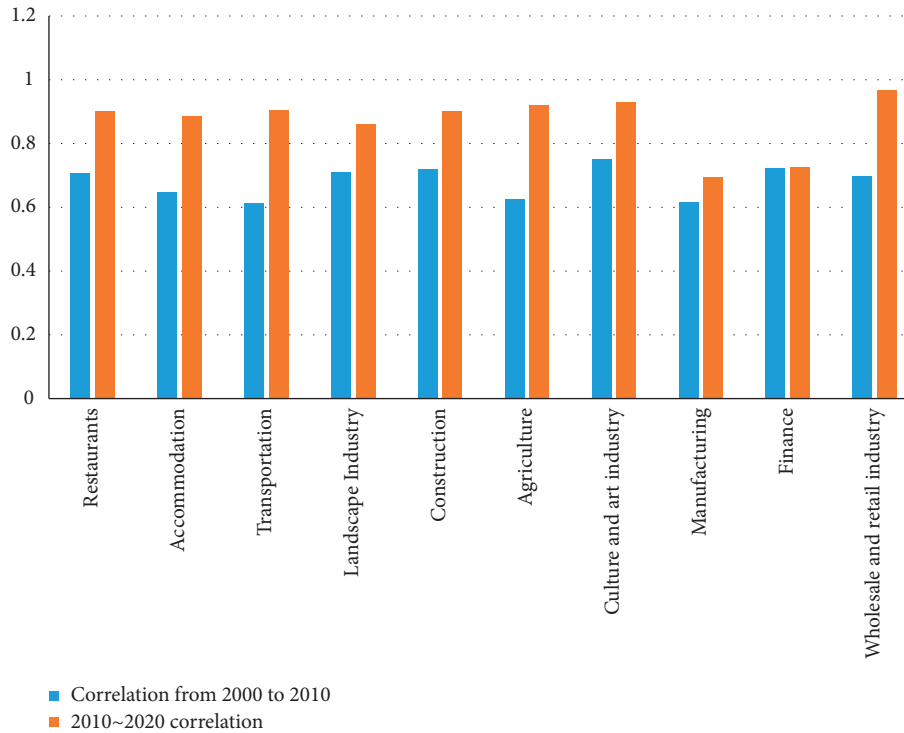


FIGURE 8: Gray correlation between the intangible cultural heritage tourism industry and related industries in two stages from 2000 to 2010 and from 2011 to 2020.

“meaning” behind the connotation and cultural lineage, is an important link between intangible cultural heritage and tourism industry resources linkage and effective synergy. The development of “ICH + tourism” products in China is an important means for cultural tourism products to achieve differentiation and gain competitive advantage, and also an inexhaustible source for cultural tourism products to further enhance their core competitiveness.

At the overall regional level, China’s “intangible cultural heritage plus tourism industry” has great economic and cultural potential; however, cultural tourism lags far behind the high level of economic development, and intangible cultural tourism resources have not been effectively developed and utilized. “The synergistic effect of “1 + 1 > 2” resources has not yet been formed. According to the evaluation of the correlation of system coupling in this study, the correlation between “intangible cultural heritage” and the tourism comprehensive development system is strong, and the foundation of the combination of the two systems is stable. Most of the intangible cultural heritage areas in China are connected to each other by mountains and rivers, with similar customs, and the cultural and tourism resources share the same root and attributes. Many areas are not only places of interest but also places where “intangible cultural heritage” projects originate and are discovered. Therefore, to strengthen the theoretical exploration of the protection of the connotation of “intangible cultural heritage + tourism” scenic resources is a major historical task for China to realize the coordinated and sustainable development of intangible oral cultural heritage and related tourism resource industries in

each region. From a perspective of promoting the combination of intangible cultural heritage inheritance and mass tourism activities, it is a study to enhance the distinctive characteristics of modern tourism cultural product development and improve the innovative quality of mass tourism consumer product development by deeply excavating the profound connotation of China’s intangible cultural traditions and infusing such intangible cultural elements into modern tourism and leisure product development. Culture and tourism consumption products market to maintain a strong sustainable development of attractive and effective two main effective way; from the ancient city tourism culture combined with the protection of non-heritage. Huaihai tourism economic zone planning should pay attention to the protection of ancient cities and towns, ecological and leisure tourism and scenic tourism development, and drive the protection of local characteristics of cultural resources tourism area of reasonable development, through the effective combination of material culture, non-material culture, and natural ecological culture, and enhance the competitiveness of regional cultural tourism.

4.3.2. Give Full Play to the “Non-Physical Heritage + Tourism” Government Guidance, and Improve the Long-Term Mechanism of Cultural Tourism Integration. Based on the evaluation of the empirical results and the benchmarking of each element index, it is necessary to give full play to the role of the government, carry out overall planning, and improve the long-term mechanism for the effective integration of “intangible heritage + tourism.” Government guidance and

improvement of mechanism are the top-level guarantee for the coordinated development of intangible cultural heritage and the tourism industry. The synergy of “ICH + tourism” in China is a systematic work involving multiple departments and fields, which requires government departments to stand at the height of the strategic overall situation, carry out scientific planning with advanced concepts, build a long-term mechanism for the integration of culture and tourism, and promote the regional “ICH + tourism” in-depth synergistic development. We should make full use of the guidance of local governments, form a coordinated and unified promotion mechanism of relevant departments, and build a development idea and long-term mechanism of “policy support, protection priority, market operation, and cultural tourism synergy.”

On the premise of the mechanism and important guarantee for the coordinated development of intangible cultural heritage and tourism industry, local governments should establish a breakthrough in the “intangible + tourism development model” in terms of institutional mechanisms, policies and standards, taking into account the specific situation of China, and provide a sustainable development direction for “intangible + tourism navigation and backup protection”. The first is to strengthen organizational leadership and improve the institutional mechanism of integrated development of culture and tourism. To establish a leading group for the development of “intangible heritage + tourism,” strengthen the style and effectiveness of government departments, to establish a sound communication and liaison mechanism, to improve the performance assessment and incentive mechanism, and to give full play to the guiding role of government departments in the development of “intangible heritage + tourism.” The second is to strengthen planning and design, and develop action plans and specific programs for the integration of cultural tourism. Adhering to the concept of advanced planning, standing in the height of an industrial layout, the development of “intangible cultural heritage + tourism” involves a number of departments and areas of overall planning and strategic deployment, the development of corresponding planning design and action plans to promote the implementation of related work. The third is to improve the policy system and establish the ICH + tourism development model. The government has introduced special policies for the integrated development of the cultural industry and the tourism industry in terms of finance, talents and land, supported the development with comprehensive policies, reasonably planned the secondary distribution of funds by cultural and tourism bureaus in various cities, provided strong guarantee for promoting the development of “intangible cultural heritage + tourism,” and formulated the “intangible cultural heritage + tourism” demonstration project selection system. The “non-foreign heritage + tourism” demonstration project selection system is formulated, and policy tilts are given to high-quality projects and high-growth enterprises to create regional “non-foreign heritage + tourism” benchmark projects and play a demonstration and a leading role [19].

4.3.3. Expand the Industrial Chain of “Non-Foreign Heritage + Tourism” in an Orderly Manner, and Promote the Sustainable Development of Cultural Tourism Synergy. Based on the analysis of the empirical research results, the key indicators and contents are aligned, and the goal is to promote the effective synergistic development of the industrial economy of the two systems with the coupled and coordinated development of “non-heritage + tourism,” which requires the extension of the industrial chain of cultural and tourism integration. The complete industrial chain shows the coordinated and healthy development of the “NRM + tourism” industry. To further promote the innovation and synergy of culture and tourism, we need to build an industrial chain as the framework support. Based on the attributes of the multifaceted linkage of “NRM + tourism,” we can effectively drive the linkage of regional industries, expand the industrial chain of “NRM + tourism” in an orderly manner, give full play to the correlation of culture and tourism industries, and promote the sustainable development of “NRM + tourism.” It is an effective path to ensure the high-quality and coordinated development of China’s intangible cultural heritage and the tourism industry [25].

In the analysis of China’s economic factors industry distribution, categories, upstream and downstream resources relationship, history and other basis, focus on judging the coordinated development of China’s economy, the status and role of the “invisible + tourism” industry, and the time, from the actual starting point and Wenxing travel, Wenxing travel, special cultural tourism, not only can extend the “invisible + travel” tourism industry chain, cultivate new tourism economic growth points but also allow more foreign tourists to understand China’s history and culture, expand the influence, attractiveness, and competitiveness of culture. Linking the local economic zone “intangible + tourism” characteristics of resources, cultural tourism around the building area performing arts, catering tourism, leisure, night market subdivision to form a full industry chain, such as ecological experience, to enhance the “intangible + tourism” travel, product quality improvement, innovation “cultural tourism industry chain+,” cultivate “non-material learning tourism,” “non-material recreation tourism” and other new economic growth points. We continuously strengthen the chain extension and realize the complementary advantages and strong combination of “intangible cultural heritage + tourism” industry chain with China’s economic zone. We insist on relying on the dual advantages of natural ecology and humanities and history, complementing and strengthening the “intangible cultural heritage + tourism” industry chain, which is a scarce link between ecological experience and historical heritage in China’s economic zone with superior natural conditions and unique regional culture, and promoting the differentiated development of cultural tourism industry. We will take the initiative to coordinate all parties, coordinate consideration, planning and promotion, deeply integrate into the cultural tourism industry chain, and create a number of “intangible cultural heritage + tourism” industry clusters with distinct advantages, high-end cutting, and chain connection.

5. Conclusion

In this paper, we refer to relevant domestic and foreign literature, and sort out the current situation of domestic and foreign research and the development of domestic intangible cultural heritage, and explain that intangible cultural heritage resources can be transformed into cultural capital and thus gain economic value through cultural capital theory. The value of intangible cultural heritage is assessed and measured by IPA model analysis, Kernel density estimation (KDE), and gray correlation method (calculated by using the results of IPA analysis). Through the analysis of the distribution quantity and density of national intangible cultural heritage and the gray correlation between the intangible cultural heritage tourism industry and related industries, the following suggestions for the development of intangible cultural heritage are proposed as follows: (1) To deeply explore the connotation of “intangible heritage + tourism” and increase the development of cultural tourism resources. To promote the coordinated development of coupling, we should improve the image of culture and tourism through deep excavation of the connotation of “intangible cultural heritage + tourism” on the basis of the integration and joint development of “intangible cultural heritage + tourism” resources. In-depth excavation of regional intangible cultural heritage and tourism common source connotation, we explore the connotation and cultural lineage behind the “reference,” which is an important link between intangible cultural heritage and tourism industry resources link, effective synergy. Product development of “intangible cultural heritage + tourism” in China is an important means of differentiating cultural tourism products and a source of core competitiveness of cultural tourism products. (2) To give full play to the guidance of the government of “non-foreign heritage + tourism” and improve the long-term mechanism for the integration of culture and tourism. It is necessary to give full play to the role of the government, carry out overall planning, and improve the long-term mechanism for the effective integration of “intangible cultural heritage + tourism.” Government guidance and improvement of mechanism are the top-level guarantee for the coordinated development of intangible cultural heritage and the tourism industry. The synergy of “ICH + tourism” in China is a systematic work involving multiple departments and fields, which requires government departments to stand at the height of the strategic overall situation, carry out scientific planning with advanced concepts, build a long-term mechanism for the integration of culture and tourism, and promote the regional “ICH + tourism.” In-depth synergistic development, we should make full use of the guidance of local governments, form a coordinated and unified promotion mechanism of relevant departments, and build a development idea and long-term mechanism of “policy support, protection priority, market operation, and cultural tourism synergy.”

Data Availability

The dataset is available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work received the analysis of red cultural resources applied to ideological and political education in colleges and universities in the new era under the Chongqing Education Commission’s Humanities and Social Sciences Research Program in 2022 (project number: 22SKSZ049).

References

- [1] J. Zhang, Z. Xie, H. Zhang, Li Yu, and D. He, “Study on the value fit and practical direction of regional intangible cultural heritage integration into aesthetic education in colleges and universities—taking yancheng’s non-legacy as an example,” *International Journal of Social Science and Education Research*, vol. 5, no. 3, pp. 585–590, 2022.
- [2] T. Fan and H. Wang, “Research of Chinese intangible cultural heritage knowledge graph construction and attribute value extraction with graph attention network,” *Information Processing & Management*, vol. 59, no. 1, Article ID 102753, 2022.
- [3] X. Ma, W. Li, and A. Bianba, “The chess theory and cultural characteristics of Chinese intangible cultural heritage “jiu” chess,” *Advances in Physical Education*, vol. 11, no. 1, pp. 82–88, 2021.
- [4] T. Lan, Z. Zheng, Di Tian, R. Zhang, R. Law, and Mu Zhang, “Resident-tourist value Co-creation in the intangible cultural heritage tourism context: the role of residents’ perception of tourism development and emotional solidarity,” *Sustainability*, vol. 13, no. 3, p. 1369, 2021.
- [5] H. Liu, “Perceived value dimension, product involvement and purchase intention for intangible cultural heritage souvenir,” *American Journal of Industrial and Business Management*, vol. 11, no. 1, 2021.
- [6] H. Zhao and X. Cao, “Analysis on the value composition of textile intangible cultural heritage,” *Scientific and Social Research*, vol. 2, no. 3, 2020.
- [7] L. Zhou, “Analysis on the protection strategy of intangible cultural heritage in contemporary China,” *Learning & Education*, vol. 9, no. 3, p. 154, 2020.
- [8] W. Wu, “On the aesthetic value and narrative path of rural intangible cultural heritage stories from the perspective of cultural education—take the legend of shun as an example,” *International Journal of Educational Technology*, vol. 1, no. 4, 2020.
- [9] Y. Lin, “On the importance of “spiritual value” in inheritance education of intangible cultural heritage,” *International Journal of Frontiers in Sociology*, vol. 2, no. 4, 2020.
- [10] H. Xiao, “Analysis of the value of intangible cultural heritage in promoting rural revitalization,” *Basic & Clinical Pharmacology and Toxicology*, vol. 127, p. 207, 2020.
- [11] X. Su, Xi Li, Y. Wu, and L. Yao, “How is intangible cultural heritage valued in the eyes of inheritors? Scale development and validation,” *Journal of Hospitality & Tourism Research*, vol. 44, no. 5, pp. 806–834, 2020.
- [12] X. Su, X. Li, Y. Wu, and L. Yao, “How is intangible cultural heritage valued in the eyes of inheritors? scale development and validation,” *Journal of Hospitality & Tourism Research*, vol. 44, no. 5, pp. 806–834, 2020.

- [13] W. U. Chen, "Analysis of intangible cultural heritage in the category of traditional oral literature," *Creativity and Innovation*, vol. 4, no. 5, 2020.
- [14] H. C. Jesús, M. L. Palma, and F. Aguado Luis, "How to measure intangible cultural heritage value? The case of flamenco in Spain," *Empirical Studies of the Arts*, vol. 39, no. 2, 2020.
- [15] X. Su, X. Li, W. Chen, and T. Zeng, "Subjective vitality, authenticity experience, and intangible cultural heritage tourism: an empirical study of the puppet show," *Journal of Travel & Tourism Marketing*, vol. 37, no. 2, pp. 258–271, 2020.
- [16] X. Kong and J. Hong, "Productive protection of intangible cultural heritage based on the theory of involvement: a case study on the farmers' paintings of Xinji county, Hebei province, China," *SHS Web of Conferences*, vol. 86, 2020.
- [17] Q. Qiu, T. Zheng, Z. Xiang, and M. Zhang, "Visiting intangible cultural heritage tourism sites: from value cognition to attitude and intention," *Sustainability*, vol. 12, no. 1, p. 132, 2019.
- [18] D. Zhou, "The value significance and protection path of the inheritance of intangible cultural heritage taking Cantonese clock design and manufacture techniques as an example," in *Proceedings of the 5th International Conference on Arts, Design and Contemporary Education (ICADCE 2019)*, Atlantis Press, August 2019.
- [19] J. A. Martilla and J. C. James, "Importance-performance analysis," *Journal of Marketing*, no. 1, pp. 77–79, 1977.
- [20] M. R. Evans and K. S. Chon, "Formulating and evaluating tourism policy using importance—performance analysis," *Hospitality Education and Research Journal*, no. 3, pp. 203–213, 1989.
- [21] Y. Geng and S. Wu, "Ethnic tourism service quality satisfaction and importance assessment and research," *Guizhou Ethnic Studies*, no. 3, pp. 171–175, 2016.
- [22] Z. Sang, F. Shuxia, P. Shi et al., "Perceptions of tourism lodging based on IPA theory: importance, satisfaction and differences--an example from Enshi Prefecture, Hubei Province," *Resource Development and Market*, no. 7, pp. 992–997, 2018.
- [23] X. Jing and F. Tan, "Research on the perception analysis of tourism destination image and government-led strategy - taking Guilin city tourism as an example," *Journal of Liaoning University (Philosophy and Social Science Edition)*, no. 6, pp. 82–90, 2018.
- [24] F. Tao, "A study on the aesthetic value and innovation of zhoushan dialect folk songs," in *Proceedings of the 2019 International Conference on Pedagogy, Communication and Sociology (ICPCS 2019)*, pp. 367–370, Atlantis Press, May 2019.
- [25] Y. Wang, "Intangible cultural landscape in Chinese agricultural cultural heritage concept, progress and value analysis," *International Journal of Social Science and Education Research*, vol. 4, no. 3, pp. 102–109, 2021.

Research Article

Study on the Innovative Development of Digital Media Art in the Context of Artificial Intelligence

Chaomiao Chen 

College of Fine Art and Design, Quanzhou Normal University, Quanzhou 362000, Fujian, China

Correspondence should be addressed to Chaomiao Chen; ccm126@qztc.edu.cn

Received 16 June 2022; Revised 16 July 2022; Accepted 19 July 2022; Published 8 August 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Chaomiao 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 modern science and technology, the speed of digital media to disseminate information is also accelerating, and the forms of communication become more and more diversified. In order to make digital media art better innovate and develop, designers should actively explore, so that art under digital media has the characteristics of intelligence, networking, and content diversification, so that it can be better applied in modern digital media communication and let visual art reach new development heights in the context of digital media. The current rapid development of artificial intelligence technology is being gradually applied to all fields of society. From the current point of view, China's digital media art will inevitably remain in the early development stage for a long time to face many difficulties and problems. In the context of the era of artificial intelligence, there will also be a large number of art and design talents for continuous research and exploration, and the emergence of these talents will certainly prompt the further development and enhancement of digital media art and artificial intelligence technology. In this process, it is especially worth noting that China's excellent cultural connotations should not be abandoned. Committed to digging more beneficial digital media art elements from traditional culture to make digital media art always develop in a more meaningful way. This paper takes the artificial intelligence era as the background, discusses the core of the development of the mutual integration of artificial intelligence technology and digital media art, analyzes the current development status of digital media art and technology, as well as the innovative development direction and future trends, proposes a digital media art design algorithm based on the convolutional neural network, and finally proves the effectiveness of the method in the relevant data set.

1. Introduction

With the increasing development of current science and technology and social economy, the art of combining digital and multimedia has gradually developed and matured, while gradually replacing traditional art methods and becoming a mainstream art. At present, it is necessary to accelerate the pace of development, reform, and innovation for the concept of art, and to fully display the characteristics of the times of the art of combining digital and multimedia. First, the need for innovation in the art of combining digital and multimedia concept, the art of combining digital and multimedia in the actual development process is still not optimistic, for example, the process of product design is seriously lacking artistry, often just applying the template to carry out design work, which is conducive to people to change the traditional cognition of the art of combining digital and multimedia

[1–3]. The art of combining digital and multimedia is a comprehensive discipline, which involves many aspects of content and knowledge; however, the art of combining digital and multimedia has not yet achieved interdisciplinary communication, but is relatively rigid in the level of artistic expression; for example, shooting a certain short film not only need to have a camera but also need to create, tuning, rendering, color mixing, etc., which involves relatively cumbersome content, so, at present, the art of combining digital and multimedia development among the change of mindset is particularly urgent. The innovative development of digital media art in the context of artificial intelligence is shown in Figure 1.

To achieve further innovation and reform of the concept of the art of combining digital and multimedia, the original traditional artistic thinking and design methods should be abandoned, and the art of combining digital and multimedia

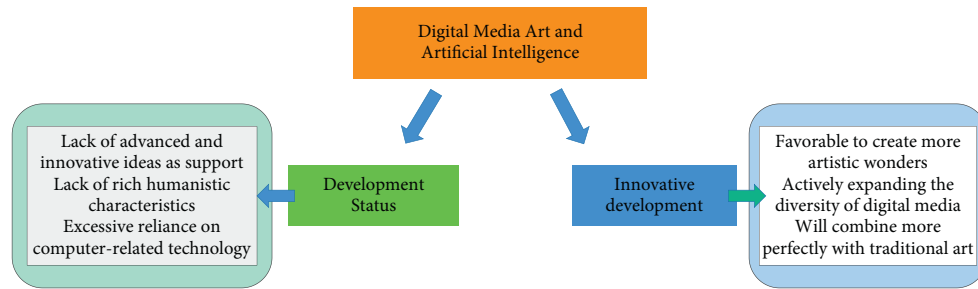


FIGURE 1: Schematic diagram of the innovative development of digital media art in the context of artificial intelligence.

with cultural connotations and contemporary values should be created [4–6]. However, thought itself has a significant dynamic character, but is also particularly susceptible to the internal and external environment, at the same time, human epiphany, association, and intuition usually have a certain influence on the thought, in order to completely out of the traditional art thought on the human fetters, need to combine with the development of the times, to cognitive and understand the new works, feel the vitality of the times, should also experience life, should only in this way can we achieve change and innovation in the art of combining digital and multimedia. In the present, artistic expression includes many types, for example, transplantation, analogy, and imagination. If traditional ideas are used in the current process of the art of combining digital and multimedia development, it is difficult to effectively meet people’s pursuit of art and requires a combination of basic artistic expression methods to promote the enrichment of the art of combining digital and multimedia connotations. First, imagination is not unrealistic but should be based on the premise of rich knowledge through the brain to effectively create and process the process. For example, if a movie is imaginative and out of touch with reality, it is unlikely to have such a great impact [7]. Although the content described in the movie is far from reality, it is still created with reality as the basis. Secondly, an analogy is mainly from the nature between two similar things, inferring things in other aspects of different points, and the same point, decision-making, by using the analogy, can often get good results. Finally, transplantation is an excellent creative technique that usually facilitates the innovation of ideas. Other techniques that are similar to transplantation include transformation, which is literally the conversion of things to each other, including conceptual and role shifts. The main content is that life is art. Nowadays, people’s material living standard has improved significantly, and the understanding of aesthetics has also undergone a great change, so art is not only exclusive to the privileged class, but everyone can create art and appreciate art, which is the test of life. Third, the expression of the art of combining digital and multimedia concept innovation is first of all the content. According to the traditional artworks for comparison and analysis, we can fully understand that the content of early artworks is relatively single. With the increasing development of the times and economy, and the mutual integration of Eastern and Western cultures, people can see more content with unique

artistic characteristics. The second is the method of creation. The rapid development of science and technology, as well as the increasing maturity of computer technology, has enabled the creation of digital media with more diverse characteristics. For example, Arlen creation software is a typical representative. Creators can practice their artistic ideas through Arlen creative software. In addition, there are still a large number of the art of combining digital and multimediatizes experimenting with other methods of art creation [8].

The world has entered the era of digital development, and the fusion of artificial intelligence technology and the art of combining digital and multimedia is rapidly reconstructing and overturning people’s perceptions of the art and design field, and new means of design and artistic expression have been created. According to the “2017 Emerging Technology Maturity Curve Report” released by a U.S. consulting firm, artificial intelligence technology had a share of over 50% of the 33 new technologies at that time [9–12]. In these subsequent years, AI technology has achieved a comprehensive and widespread application. Nowadays, AI technology is compatible with intelligent systems such as environment perception, memory storage, thinking and reasoning, and learning ability possessed by human intelligence. And in the field of design, AI-aided design has made certain achievements, such as the artificial intelligence product Laban launched by Alibaba in the Double Eleven, which designed 4×10^8 advertising banners by itself, witnessing the powerful power of art and technology collision. However, in the process of artificial intelligence replacing the basic design of artwork, people also perceive the lack of judgment and understanding of aesthetics by intelligent machines, and the beauty of art cannot be realized by computer technology alone. At this point, art designers became “trainers,” they need to cultivate the computer’s sense of beauty so that the computer can design and learn to acquire a kind of human-based aesthetics and wisdom, so as to form their own thinking and insights on design aesthetics, and then feedback to the works involved. The optimists believe that the future of the art of combining digital and multimedia has unlimited possibilities, while the pessimists believe that it will be the end of the road for art designers. Along with the accelerated pace of global digital development, Chinese digital media design majors should take the initiative to accept new knowledge and technology in education, objectively and rationally view the opportunities

and challenges brought by artificial intelligence and its high technology, and always uphold the principles of integrating tradition and modernity, approaching ideals and reality, giving equal importance to practice and theory, and carrying out reform and survival, so as to build a digital media design teaching system that meets the modern context. This paper proposes a system based on the convolutional neural network. The main contributions of this paper are (1) A convolutional neural network-based art design algorithm for combining digital and multimedia is proposed, the framework structure of the neural network model is analyzed, and the process of how to use the model to extract artistic styles and fuse them with ordinary images is discussed. (2) Then, based on the theory, according to the actual characteristics of the art combining digital and multimedia, the experimental analysis is performed to find the processing content images with suitable convolution layers, as well as finding the best overlay combination for digital media art feature extraction and proposing visualization criteria for evaluating image quality. (3) Finally, the feasibility of the theory is verified by adjusting the scale coefficients of content images and style images to obtain images that meet the expected goals, and a new style extraction method for digital and multimedia combined art is proposed.

2. Related Work

2.1. Digital Media Art. The art of combining digital and multimedia refers to the synthesis of digital image processing technology, information and communication technology, art design, and other disciplines. Unlike traditional art, the art of combining digital and multimedia is widely used and developed in other fields through media communication with its own unique characteristics of integration, virtualization, and intersectionality [12–14]. The art of combining digital and multimedia has diversified forms of expression, and it changes people's aesthetic pursuit of art with its unique creation concept, influencing the direction of art modernity development. As socialism with Chinese characteristics enters a new era, the living standards of the people and the aesthetic ability of the public are increasingly rising, the art of combining digital and multimedia not only brings people a better quality visual experience but also gives them stronger sensory stimulation. In artistic expression and communication, the art of combining digital and multimedia can vividly and effectively release and convey information and is gradually becoming a new carrier for the development of modern artistic expression. The benign development of the art of combining digital and multimedia can play an excellent role in promoting the development of human society. The main characteristics of the art of combining digital and multimedia along with the development of social life, people have put forward higher requirements for the accuracy, interactivity, and efficiency of information dissemination, and the application of the art of combining digital and multimedia in art has brought about great changes in people's artistic feelings. The art of combining digital and multimedia is an art that combines participation, interactivity, multimedia use, high-tech

enrichment, and new forms of expression, and is a reorganization of the real world and virtual images. Its emergence and development has provided a broader stage of expression for traditional art content. Grassroots participation in art is closely related to people's lives, and people can enjoy their favorite artworks at any time, publish their own comments on artworks through the Internet, and even publish their own artworks anonymously. 3D and 4D technologies that were only bred in the digital media era, have become popular, extending people's ability to appreciate art, innovating forms of art appreciation, and getting closer to art. The interactive nature of two-way interactive is the art of combining digital and multimedia.

The interactive nature of two-way interactive is the art of combining digital and multimedia provides a new creative experience for art creators. Before the digital media came out, the dissemination of art was a whole process of creation, release, and transmission by artists and traditional media, with the audience only playing the role of receiver and appreciator, in which the artist could not get feedback from the audience. The emergence of digital media has brought about a new change in the process of art communication [15–17]. The whole process has changed from one-way to two-way interaction, and the artist can also know the feelings of the audience at any time, and the audience can also participate in the creation of the process with two-way interaction. Multimedia use of our country's traditional artworks are created through different tools and different materials, for now, China's art of combining digital and multimedia products are mainly conducted through computers and new technologies. In the process of painting oil paintings, the creators first use computer software, which will help with later editing and modifications, and the computer can record the entire creative process of the creator. In addition to these, the art of combining digital and multimedia also includes e-books and e-maps. People use e-books for reading and can make comments and changes at any time, bringing convenience to reading. People can enter the places they want to go on the electronic map, and it will remind people of the relevant traffic precautions. One of the most important features of high-tech enrichment in the field of art is exaggeration, and the art of combining digital and multimedia is no exception, and it brings this feature to the highest point. We can easily find that in the art of combining digital and multimedia, every piece of art and every character is exaggerated, with exaggerated expressions and shapes, and even exaggerated language, all of which are designed to give people the best artistic experience. For example, new technologies in the art of combining digital and multimedia are used in paintings to make them more vivid and rich; VR games are games with high technology in the art of combining digital and multimedia, and the game screen becomes three-dimensional from two-dimensional, so that people can participate in the game more realistically. Compared with other types of artistic expression, the art of combining digital and multimedia has its own unique innovative performance, which is summarized in the following points: innovation of creative concepts China's economy in the new era is still in a relatively rapid development period, and the strong

economic level provides convenient conditions for the innovative development of the art of combining digital and multimedia. The economic foundation determines the superstructure, with the economic foundation as a pavement, and then through new technology and means, the art of combining digital and multimedia can have adequate development concepts and new ideas. Under the new creation concept, the traditional means and the current means will be very different, the creation method is more and more rich and diverse, the creation content is newer, the art of combining digital and multimedia also pays more attention to the development of personality, and the needs of different audience groups are satisfied. The difficulty of creation lies in new ideas, once any creation lacks new ideas, it is difficult to get people's love, and the art of combining digital and multimedia is no exception [17–19]. With the emergence of the art of combining digital and multimedia, some old artists do not have a deep enough understanding of this new form of creation and cannot break through themselves in time with the development of the times. Some people still use the old methods in their creation and simply think about the development of traditional art according to their past thinking. In their eyes, the modern art of combining digital and multimedia has no way to reflect the real value of their works. This makes the works they create lack their own individuality and are as same. Besides these people, there are also some artists who are not too familiar with the functions of digital media, leading to the lack of innovation in their works and the phenomenon of similar works.

2.2. Artificial Intelligence Technology. The integration of artificial intelligence technology and the art of combining digital and multimedia is based on human understanding of themselves and their experience of emotions [19–21]. If a person has no understanding of the formation of aesthetic thinking and consciousness in a different cultural context, it is simply impossible to solve various design problems concerning the way humans live and work if they simply rely on the computational methods of intelligent machines. Today's iterative developments in science and technology continue to enrich the way people live and produce and raise the demand for aesthetics. The reason for developing artificial intelligence technology is to better serve people themselves, while the integration of art and design is precisely from the emotional needs of people, where art is more prominent in the expression of the subjective concept of people. The future of artificial intelligence is to think about the needs of people and the ability to change different environmental needs. From the current application of artificial intelligence scenarios, artificial intelligence in solving certain boundary problems, people are bound to lose to the machine; however, the need to make decisions with the help of emotional judgment, the logical reasoning made based on artificial intelligence big data also exists uncertainty. At present, artificial intelligence can only provide high technology, while the formation of excellent artworks must require new technical means to achieve. People have emphasized that art is the expression of human subjective

concepts, just like saying that if it is similar, it is vulgar, and if it is not, it is not. Art does not have boundaries; therefore, artificial intelligence can never replace a human to complete subjective concepts. The key to the integration of artificial intelligence technology and the art of combining digital and multimedia is to obtain new methods of art created with the help of artificial intelligence technology, so as to better enrich the means of the art design, so the core of the integration of the two depends on people's values and cognitive ability of artificial intelligence [22].

In summary, we know the concept of artificial intelligence, the way of integration of artificial intelligence and the art of combining digital and multimedia and the core significance. In the context of artificial intelligence, intelligent machines can drive the continuous development and progress of society by virtue of their powerful computing power, and the next stage of deeper integration and innovation of artificial intelligence and the art of combining digital and multimedia will definitely present the following development characteristics. The art of combining digital and multimedia has characteristics such as openness, integration, and interactivity, and there are also artistic time-sensitive features that change with time or forms, and it is based on these special features that make the current art of combining digital and multimedia more and more diversified. In recent years, the rapid development of computer and network technology in China has reached the needs of art creation making the organic integration of art and artificial intelligence technology. Only by scientifically and rationally utilizing this fusion can the art of combining digital and multimedia give audiences a more perfect visual experience and strengthen their perceptions and impressions of the works of art creation. With various high-tech as the creative support of the art of combining digital and multimedia, it will let art and technology burst into more artistic miracles. Actively expanding the diversity of digital media, the characteristics of the art of combining digital and multimedia are more diverse, yet in fact, the development of the art of combining digital and multimedia and artificial intelligence technology shows an obvious sense of weakness. In this period of sustainable development of the art of combining digital and multimedia and simultaneous technological progress, Chinese art of combining digital and multimedia will present more new types of art in a richer form of expression, prompting digital media to achieve diverse development, bringing audiences a completely different consumer experience and a new experience and artistic perceptiveness in visual perception [23–25]. From this point of view, the mutual integration of the art of combining digital and multimedia and artificial intelligence technology will certainly make the development space of art and technology more extensive. From the analysis of the development history of the art of combining digital and multimedia in the past, the integration of the art of combining digital and multimedia and artificial intelligence technology in China can only be better presented in the form of technology and art through digital media by continuously strengthening its own traditional cultural heritage and modern civilization. Most people one-sidedly believe that

digital technology and computer technology is the catalyst for the development of the art of combining digital and multimedia and technology, however, in fact from another perspective, it is based on the background of the artificial intelligence era that society has stepped into a new period of development of art and technology. From the current point of view, Chinese art of combining digital and multimedia is bound to remain in the early stages of development for a long time and to face many difficulties and problems. In the context of the artificial intelligence era, there will also be a large number of art and design talents emerging for continuous research and exploration, and the emergence of these talents will certainly lead to further development and enhancement of the art of combining digital and multimedia and artificial intelligence technology. In this process, it is especially worth noting that the excellent Chinese cultural connotations should not be abandoned. Commitment to digging out more beneficial art of combining digital and multimedia elements from traditional culture is the only way to make the art of combining digital and multimedia always develop in a more meaningful way.

3. Methods

3.1. Model Architecture. The art of combining digital and multimedia style conversion is an important technique for nonrealistic drawing in computer graphics. For texture synthesis, existing nonparametric algorithms can synthesize realistic natural textures by resampling the pixels of a given source texture. Most previous texture transfer algorithms use these nonparametric methods for texture synthesis, while using different methods to preserve the structure of the target image. This study uses a deep convolutional neural network to learn a generic feature representation that performs texture transfer (i.e., style transformation) while preserving the semantic content of the target image. The model structure of this paper is shown in Figure 2.

3.2. Image Semantic Content Representation. Given an input image x , each layer of the convolutional neural network is encoded in each layer using the image's filter. A layer of the neural network contains N_l different filters, i.e., it has N_l feature mappings, each of size M_l . Therefore, the response in layer l can be stored in the matrix F^l , where F_{ij}^l is the activation function of the i th filter at position j in layer l . To visualize the image information encoded on different layers, gradient descent can be performed on the white noise image to find another image that matches the feature response of the original image. Let p and x be the original image and the generated image, and let P^l and F^l represent the features in the l th layer, respectively. Then, define the squared error loss between the two feature representations as follows:

$$L_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - P_{ij}^l)^2. \quad (1)$$

With respect to the activation function in layer l , the derivative of the loss function as

$$\frac{dL_{\text{content}}}{dF_{ij}^l} = \begin{cases} (F^l - P^l)_{ij}, & \text{if } F_{ij}^l > 0, \\ 0, & \text{if } F_{ij}^l < 0. \end{cases} \quad (2)$$

The standard error backpropagation can then be used to calculate the gradient with respect to the image x . Thus, the initial random image x can be changed until it generates the same response in a particular layer of the convolutional neural network as the original image p . The schematic diagram of semantic content representation and transformation is shown in Figure 3.

3.3. Artistic Style Representation. The generative network model in the art of combining digital and multimedia style network uses a residual network, which uses a "bootstrap" approach to effectively solve the problem of gradient disappearance when the deep learning network is too deep. In addition, the advantage of using the residual layer to build the network is that the residual layer has a faster training speed than the general convolutional layer with the same convolutional effect. The style transition diagram is shown in Figure 4. The perceptual network model uses a five-layer residual network for image feature extraction; however, the perceptual network does not have a deep depth, with three convolutional layers, five residual layers plus three deconvolutional layers for a total of eleven layers. Each residual layer contains two convolutional layers and the size of the convolutional kernel is 3×3 . The art of combining digital and multimedia generation network is deepened and transformed sufficiently based on this, and ten residual layers are chosen for the middle part of the art of combining digital and multimedia generation network to play the function of the residual network as much as possible under the condition of the existing experimental environment. At the same time, the residual layer of the art of combining digital and multimedia generation network is adjusted from two 3×3 convolutional kernels to a combination of two 1×1 convolutional kernels and one 3×3 convolutional kernel. 1×1 convolutional kernels are not useful in two-dimensional plane operations, but each network layer of the art of combining digital and multimedia generation network has multiple channels for receiving image data and its response, and each network layer has multiple convolutional kernels. At this point, the 1×1 convolutional kernel can downscale these stacked 3D data so that the number of input and output channels of the convolutional layer is reduced, thus reducing the number of network parameters, and the training speed of the network model is significantly improved. Considered from the opposite perspective, due to the use of 1×1 convolutional kernels, the art of combining digital and multimedia generation network can be allowed to receive image data with larger dimensions at the same network size. For example, the original art of combining digital and multimedia generation network allows input images with a maximum dimension of 512×512 , while the improved network can accommodate images with a resolution of 2048×2048 . The selection of better quality and larger resolution images will result in better visualization of the experimental results. These feature correlations are represented by the Gram matrix G_{ij}^l , where G_{ij}^l is the inner product between vectorized feature mappings i and j in layer l .

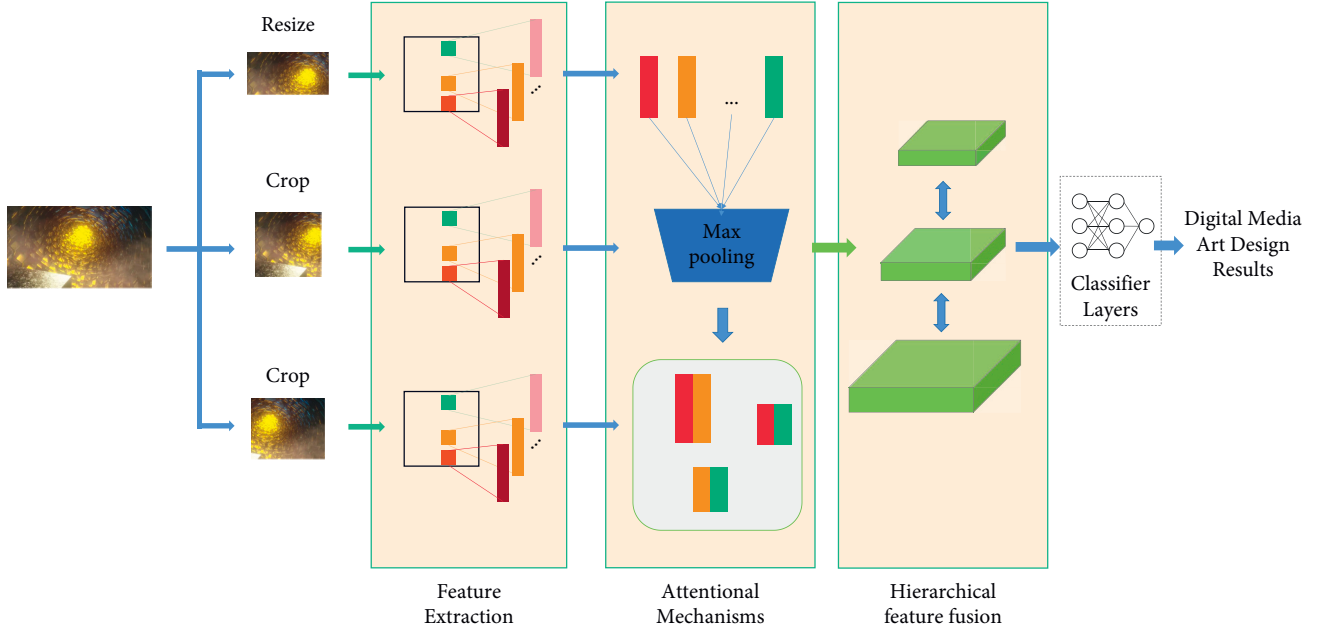


FIGURE 2: Model structure.

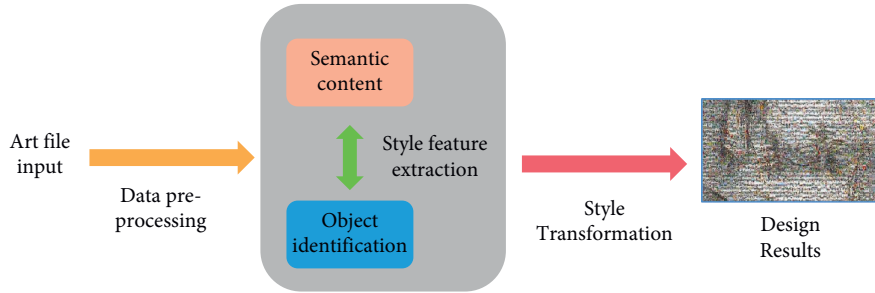


FIGURE 3: Schematic diagram of semantic content representation and transformation.

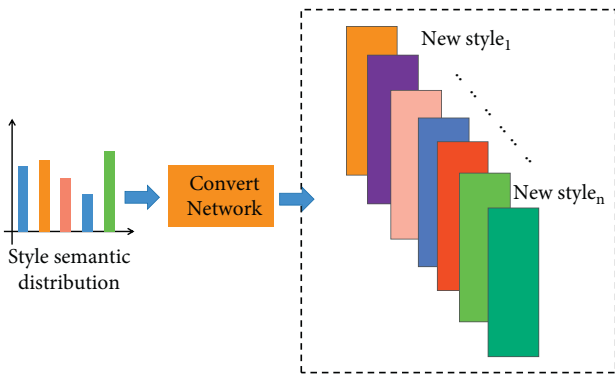


FIGURE 4: Schematic diagram of style transformation.

$$G_{ij}^l = \sum_k F_{ik}^l F_{jk}^l. \quad (3)$$

Let a and x be the original image and the generated image, while A_{ij}^l and G_{ij}^l denote the styles of the original and generated image layers l , respectively. Then, the loss of layer l can be expressed as

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2. \quad (4)$$

The total loss function has the following formula:

$$L_{\text{style}}(\vec{a}, \vec{x}) = \sum_l w_l E_l, \quad (5)$$

where w_l is the weighting factor. With respect to the activation function in layer l , the derivative of E_l can be expressed as

$$\frac{\partial E_l}{\partial F_{ij}^l} = \begin{cases} \frac{1}{N_l^2 M_l^2} \left((F^l)^T (G^l - A^l) \right)_{ji}, & \text{if } F_{ij}^l > 0, \\ 0, & \text{if } F_{ij}^l < 0, \end{cases} \quad (6)$$

The gradient of E_l with respect to the pixel value x can be easily calculated using standard error back propagation.

3.4. Training the Art of Combining Digital and Multimedia Style Network. The art of combining digital and multimedia

style network is constructed, and similar to other CNN networks, the network needs to rely on a large amount of image data and loss functions for training. In this paper, we use the MSCOCO2014 dataset, which contains 328,000 images and is divided into 91 categories. The update of the art of combining digital and multimedia generation network parameters relies on the loss function and the stochastic gradient descent algorithm, and the specific process of training is as follows: firstly, the noisy image x is input into the art of combining digital and multimedia generation network, and x is convolved and deconvoluted by the network operation to obtain images of the same size as x . The content map y and a specified art of combining digital and multimedia image y in the MSCOCO dataset are used in the art of combining digital and multimedia discriminative, the calculation results are fed back to the art of combining digital and multimedia generation network using the BP algorithm, and the parameters and weights of each layer of the art of combining digital and multimedia generation network are updated to minimize the total loss cost. Since the dataset provided to the art of combining digital and multimedia generation network for training is very large, the network has a good generalization ability after the training is completed, and it is able to apply the style of the art of combining digital and multimedia to any unknown picture. In this paper, content loss function $\ell_{\text{feat}}^{\phi,j}(\hat{y}, y_c)$ and style loss function $\ell_{\text{style}}^{\phi,j}(\hat{y}, y_s)$ are constructed for picture content and the art of combining digital and multimedia style, respectively, and the content loss function calculates and y . The square of the Euclidean distance between the response \hat{y} and y_c obtained at a layer j after input pretraining the model, where $C_j H_j W_j$ refers to the data obtained at layer j after computing the size and dimensionality of the feature mapping.

$$\ell_{\text{feat}}^{\phi,j}(\hat{y}, y_c) = \frac{1}{C_j H_j W_j} \phi_j(\hat{y}) - \phi_j(y_c)_2^2. \quad (7)$$

The extraction of image styles relies on the Gram matrix, Gram matrix $G_j^{\phi}(x)_{c,c'}$ calculates the inner product between the same dimensional feature mapping $\phi_j(x)_{h,w,c}$ and $\phi_j(x)_{h,w,c'}$ obtained at layer j . The covariance operation in the Gram matrix can be used to understand the covariance relationship between different feature mapping of responses in the same layer in the network, and also to know which network nodes are activated, responsive and cooperative at the same time. The style loss function is defined as the F -parametric of the difference between $G_j^{\phi}(\hat{y})$ and $G_j^{\phi}(y_s)$.

$$G_j^{\phi}(x)_{c,c'} = \frac{1}{c_j H_j W_j} \sum_{h=1}^{H_j} \sum_{w=1}^{W_j} \phi_j(x)_{h,w,c} \phi_j(x)_{h,w,c'}, \quad (8)$$

$$\ell_{\text{style}}^{\phi,j}(\hat{y}, y_s) = \frac{1}{c_j H_j W_j} G_j^{\phi}(\hat{y}) - G_j^{\phi}(y_s)_F^2.$$

Besides, this paper uses the full variational regularization $\lambda_{TV} \mathcal{L}_{TV}(\hat{y})$ to ensure the smoothness of the final generated images. The weighted sum of the content loss function $\lambda_c \ell_{\text{feat}}^{\phi,j}(\hat{y}, y)$, the style loss function $\ell_{\text{style}}^{\phi,j}(\hat{y}, y)$, and the full-

variance regularization $\mathcal{L}_{TV}(\hat{y})$ yields the final global loss function.

$$\mathcal{L}_{\text{total}} = \arg \min_y \lambda_c \ell_{\text{feat}}^{\phi,j}(\hat{y}, y) + \lambda_s \ell_{\text{style}}^{\phi,j}(\hat{y}, y) + \lambda_{TV} \mathcal{L}_{TV}(\hat{y}). \quad (9)$$

The training goal of the art of combining digital and multimedia generation network is to minimize the global loss function and use this value to update the response while using the BP algorithm back to the neurons of the art of combining digital and multimedia generation network, which use the gradient descent algorithm to update the parameter weights. The whole process of training the art of combining digital and multimedia generation network can also be regarded as using the feature extraction capability of the VGG-19 model parameters to help the art of combining digital and multimedia generation network obtain the corresponding art of combining digital and multimedia style simulation parameters based on the existing VGG-19 pretraining model. The trained art of combining digital and multimedia generation network has a strong generalization ability and can map the art of combining digital and multimedia style to any unknown image. Since the parameters of the generative network are mature enough, the next image that needs to map the art of combining digital and multimedia style does not need to calculate the loss function, and a forward propagation operation can be performed directly in the art of combining digital and multimedia generative network, so the art of combining digital and multimedia generative network is very fast in style mapping.

4. Experiments and Results

4.1. Experimental Setup. In this paper, we trained and used the art of combining digital and multimedia style network model on a computer with Ubuntu 16.04, GTX1070 graphics card, 8 GB RAM, and TensorFlow as the experimental framework are continuously optimized. This chapter will focus on showing the generated images using different art of combining digital and multimedia style models for stylized mapping, where the art of combining digital and multimedia style network models include the art of combining digital and multimedia models of the slate category with rough textures and bright colors and the art of combining digital and multimedia models of the paper category with delicate brush strokes and a bias toward realism, and the content images include five categories such as people, architecture, landscape, objects, and animals. These perspectives were used to view the degree of similarity between the generated images of the art of combining digital and multimedia style network and real the art of combining digital and multimedia. A total of 25 art of combining digital and multimedia style network models were trained for comparative analysis with many different categories of real the art of combining digital and multimedia. The real art of combining digital and multimedia can be broadly classified into board the art of combining digital and multimedia and paper the art of combining digital and multimedia according to the drawing method and visual effect. Before training the network

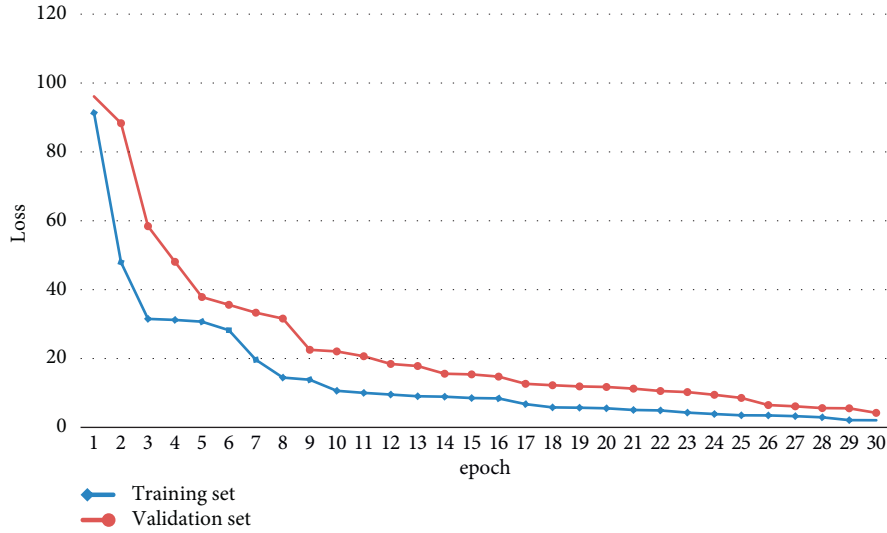


FIGURE 5: Training process loss convergence curve.

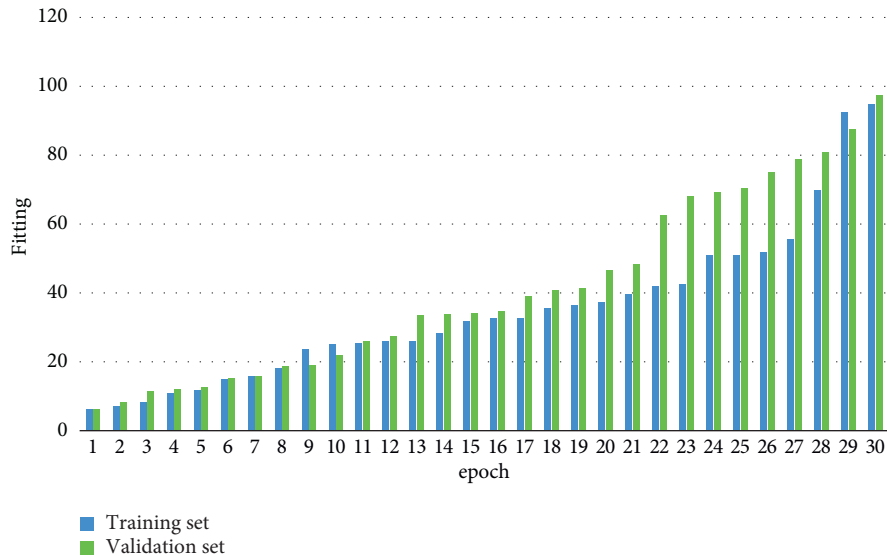


FIGURE 6: Training process performance improvement diagram.

models, different training parameters were set based on the different styles of real the art of combining digital and multimedia. One of the main parameters is the ratio of content weights to style weights in the loss function, i.e., the ratio of the weights of the content loss function to the weights of the style loss function α/β , the higher the ratio, the less stylized the generated images are, and vice versa. For the board-like the art of combining digital and multimedia network model, the value of α/β should be set larger in order to retain the roughness of lines and the visual effect of color interweaving in the real board-like the art of combining digital and multimedia as much as possible, and 0.15 is chosen as the final parameter in this paper. The value of α/β should be smaller, and 0.3 is chosen as the final parameter. The training process loss convergence curve and performance improvement are shown in Figures 5 and 6.

TABLE 1: Classification results before and after data enhancement.

Image data processing	Accuracy rate
Without data enhancement	77.34
With data enhancement	86.55

4.2. Experimental Results. Using the network model for the experiments on the original data and the data after data enhancement, the results are shown in Table 1, and it can be seen that the classification accuracy after data enhancement is improved by 9.21%.

Comparison of the results of different classification methods: comparison with the traditional network model. In order to verify the effectiveness of this paper's network model for classifying the art of combining digital and multimedia images, the data of this paper were input into the

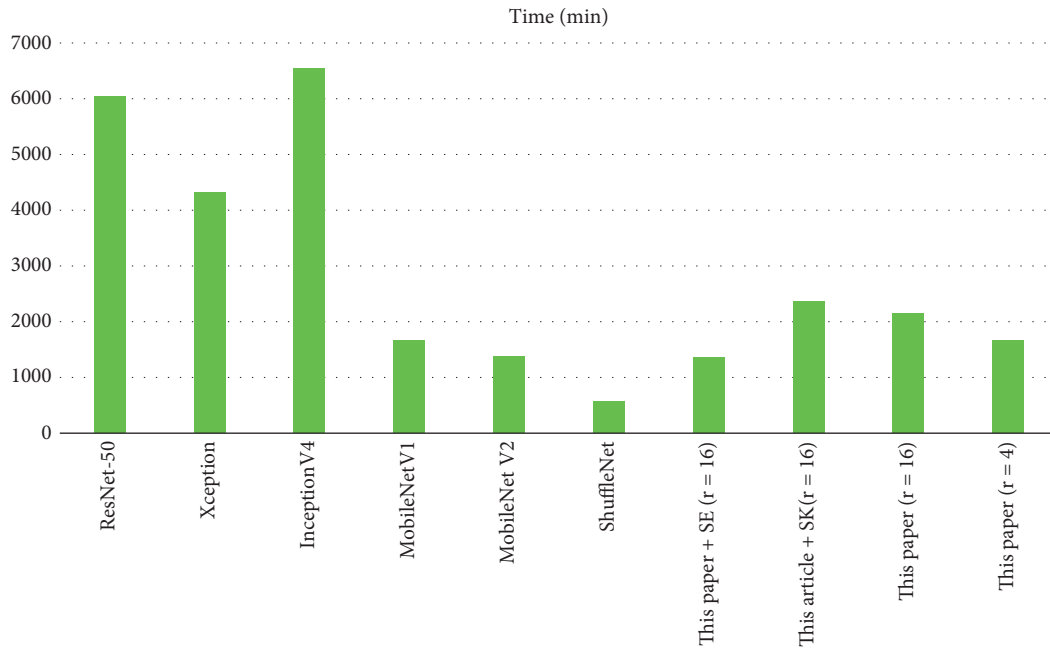


FIGURE 7: Comparison of classification time of different network models.

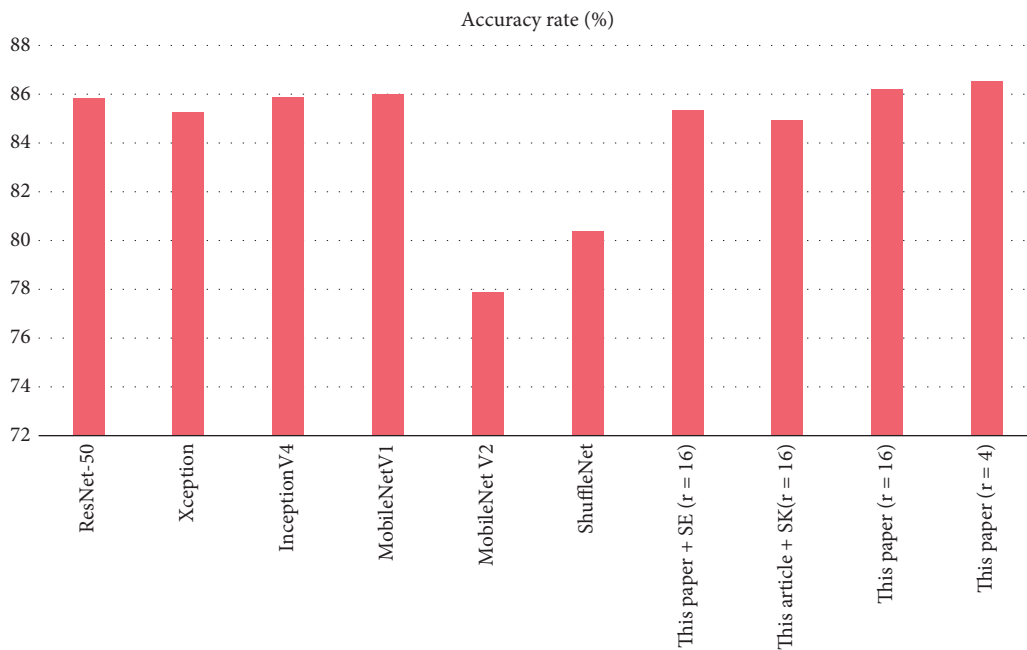


FIGURE 8: Comparison of the classification results of different network models.

traditional network model and this paper's network model for training and verification, and the results are shown in Figures 7 and 8.

From Figures 7 and 8, compared with the existed network model, the network model in this paper is lightweight, has short training time, and has the best accuracy rate. However, the proposed network module contains parallel convolution operations, which makes the parameters of the network model in this paper higher than those of Shuffle-Net, MobileNetV1, and MobileNetV2. The method + SE and the method + SK in

this paper are network models that replace the proposed modules with the SE and SK modules, and the values of the descent rate r in the modules are taken as 16. The experimental results show that the accuracy of the network model in this paper is 0.86% and 1.25% higher than that of the method + SE and the method + SK in this paper, respectively. When the r value of the proposed module is taken as 4, the classification accuracy is higher than that when r is taken as 16.

Comparing with the traditional methods, it can be seen from Table 2 that the traditional manually extracted

TABLE 2: Comparison of the results of the method in this paper with the traditional method.

Methods	Accuracy rate (%)
Comparison method 1	66.78
Comparison method 2	60.2
Method of this paper	86.55

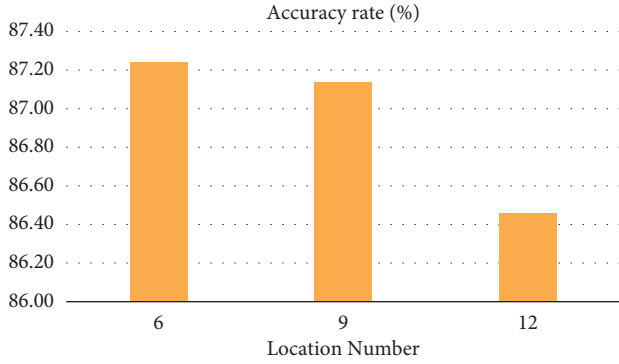


FIGURE 9: Accuracy of the module in different positions.

underlying overall features based on color, shape, etc., and local features cannot fully distinguish the style features of various types of art images, while the method in this paper can better extract the overall features and local detail features of art images and improve the classification accuracy.

Module role: location is to see the effect of the proposed module on the classification of art images at different locations of the network model, it is placed separately at the locations of the network model numbered as 6, 9, and 12 in this paper, and the convolutional kernels of 3×3 and 5×5 are taken on the branches, respectively, and the descent rate r value is taken as 4. The results are shown in Figures 9 and 10.

From Figures 9 and 10, the proposed module has the highest classification accuracy and the least computational consumption for the art images when it is placed at the position of network model number 6 alone.

The drop rate r and the branch convolution kernel size are an important set of parameters in the proposed module for controlling the computational resources and experimental accuracy. The proposed module is individually placed in the network model with numbering position 6, and the proposed r value and the branch convolution kernel size are experimented and analyzed, and the results are shown in Table 3. It can be seen that when the size of the convolution kernel on the branch of the proposed module is fixed, the classification results are higher when the r value is taken as 4 than when the r value is taken as 16; when the r value is fixed, the proposed module of 2 branches takes less time and fewer parameters than the training of 3 branches; when the convolution kernel on the branch of the proposed module is taken as 1×1 and 5×5 , respectively, the experimental results have the highest accuracy rate.

Null convolution: in order to compare the effect of null convolution kernels on art image feature extraction, different sizes of null convolution kernels are taken on the proposed module for experiments on the data of this paper,

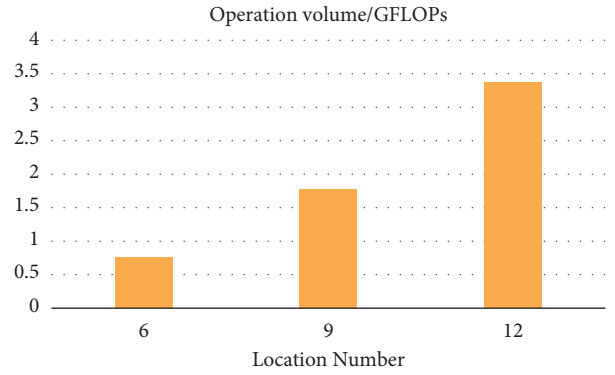


FIGURE 10: Operation volume of the module in different positions.

TABLE 3: Comparison of drop rate and convolution kernel size classification results.

Decline rate	1×1	3×3	5×5	Parameter (M)	Time (min)	Accuracy rate (%)
$R = 4$	✓	✓		2.3	1367	87.26
	✓		✓	2.6	1680	87.58
		✓	✓	2.7	1800	87.24
	✓	✓	✓	2.7	2370	87.35
$R = 16$	✓	✓		2.3	1230	86.35
	✓		✓	2.6	1220	86.85
		✓	✓	2.7	1440	86.15
	✓	✓	✓	2.7	1770	86.42

TABLE 4: Module branching takes the classification results of different null convolution kernels.

Decline rate	$K = 3$	$K = 5$	$K = 7$	Parameter (M)	Time (min)	Accuracy rate (%)
$R = 4$	✓	✓		2.4	1530	86.07
	✓		✓	2.4	1657	86.00
		✓	✓	2.4	1560	84.50
	✓	✓	✓	2.6	2220	86.40

and the results are shown in Table 4, where $K3$ denotes an ordinary 3×3 convolution kernel, $K5$ denotes a 3×3 convolution kernel with a null convolution expansion rate of 2 and a perceptual field of 5×5 , and $K7$ denotes a 3×3 convolution kernel with an expansion rate of 3 and a perceptual field of 7×7 .

The experimental results show that the parameters of the null convolution are less than those of the normal convolution with the same receptive field, but the classification accuracy is not as high as that of the normal convolution. The 3×3 convolutional kernels with 5×5 receptive fields and $K7$ indicates the 3×3 convolutional kernel with 3 expansion rate and 7×7 receptive fields.

5. Conclusion

Along with the accelerated development of artificial intelligence technology, designers in the traditional design field will be gradually outlawed by intelligent machines if they do not review the situation, transform, and improve themselves

in time. However, not all the work in the field of art and design can be replaced by it, especially some creative design concepts and contents that no high-tech and machine can do. As a professional, in the field of art and design, we should think more about the intangible value in art and design and only the core competitiveness of design talents is to burst out a constant source of creativity and inventiveness. The powerful computing power of artificial intelligence can assist them to use historical big data to explore human needs comprehensively, help people to solve practical problems through design, and use data to obtain design thinking. At the same time, the current global information is extremely rich and diverse, and this information will drive people to accept new products, new technologies, and new digital life in a passive mode. In-depth exploration of the aesthetics and wisdom of digital media design requires art designers to look at the changing times with their own understanding and feel the difference in the life of the times. This paper mainly proposes a method based on the convolutional neural network to better achieve feature extraction and classification of art images. This paper solves the problem of insufficient research on the classification of existing multi-class art images, and achieves better classification results of art images than existing network models and traditional classification methods. In the subsequent work, we will optimize the art image classification network model, expand the art image sample library, and further improve the accuracy of art image classification and the efficiency of network model classification.

Data Availability

The datasets used during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] S. Ding, S. Qu, Y. Xi, and S. Wan, "A long video caption generation algorithm for big video data retrieval," *Future Generation Computer Systems*, vol. 93, pp. 583–595, 2019.
- [2] Z. Gao, Y. Li, and S. Wan, "Exploring deep learning for view-based 3D model retrieval," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 16, no. 1, pp. 1–21, 2020.
- [3] Z. Gao, H. Xue, and S. Wan, "Multiple Discrimination and Pairwise CNN for view-based 3D object retrieval," *Neural Networks*, vol. 125, pp. 290–302, 2020.
- [4] S. Ding, S. Qu, Y. Xi, and S. Wan, "Stimulus-driven and concept-driven analysis for image caption generation," *Neurocomputing*, vol. 398, pp. 520–530, 2020.
- [5] J. Pei, Z. Yu, J. Li, M. A. Lakshmana, and K. Lakshmana, "TKAGFL: a federated communication framework under data heterogeneity," *IEEE Transactions on Network Science and Engineering*, p. 1, 2022.
- [6] Z. Niu, S. Xiang, and M. Zhang, "Application of artificial intelligence combined with three-dimensional digital technology in the design of complex works of art[J]," *Wireless Communications and Mobile Computing*, p. 2022, 2022.
- [7] S. A. Al Hashimi, A. A. Al Muwali, Y. E. Zaki, and N. A. Mahdi, "The effectiveness of social media and multimedia-based pedagogy in enhancing creativity among art, design, and digital media students," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 14, no. 21, pp. 176–190, 2019.
- [8] B. Xu and J. Jiang, "Exploitation for multimedia asian information processing and artificial intelligence-based art design and teaching in colleges," *Transactions on Asian and Low-Resource Language Information Processing*, 2022.
- [9] J. Wu, X. Wang, Y. Dang, and Z. Lv, "Digital twins and artificial intelligence in transportation infrastructure: classification, application, and future research directions," *Computers & Electrical Engineering*, vol. 101, Article ID 107983, 2022.
- [10] S. Karnouskos, "Artificial intelligence in digital media: The Era of deepfakes," *IEEE Transactions on Technology and Society*, vol. 1, no. 3, pp. 138–147, 2020.
- [11] R. Beaumont, "AI as DJ: artificial intelligence as the arbiter of musical taste in digital media platforms and the cultural continuum," *New Explorations: Studies in Culture and Communication*, vol. 1, no. 2, 2020.
- [12] N. Hong, "Digital-media-based interaction and dissemination of traditional culture integrating using social media data analytics," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5846451, 11 pages, 2022.
- [13] V. S. Viswanathan, P. Toro, and G. Corredor, ".,," *The Journal of Pathology*, vol. 257, 2022.
- [14] J. Zhao, L. Guo, and Y. Li, "Application of digital twin combined with artificial intelligence and 5G technology in the art design of digital museums," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 8214514, 12 pages, 2022.
- [15] C. Wu, "Application of digital image based on machine learning in media art design," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8546987, 14 pages, 2021.
- [16] R. Eugeni and P. Pisters, "The Artificial Intelligence of a Machine: Moving Images in the Age of algorithms," *artificial intelligence*, 2020.
- [17] Y. Cao, Z. Han, R. Kong, C. Zhang, and Q. Xie, "Technical composition and creation of interactive installation art works under the background of artificial intelligence," *Mathematical Problems in Engineering*, vol. 2021, Article ID 7227416, 11 pages, 2021.
- [18] J. W. Hong and N. M. Curran, "Artificial intelligence, artists, and art," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 15, no. 2s, pp. 1–16, 2019.
- [19] P. Galanter, "Artificial intelligence and problems in generative art theory," *Proceedings of EVA London*, vol. 2019, pp. 112–118, 2019.
- [20] D. R. Wheeler, "Art, artificial intelligence, and aesthetics in plastic surgery," *Plastic and Reconstructive Surgery*, vol. 148, no. 3, pp. 529e–530e, 2021.
- [21] C. Lin, "Application of traditional cultural symbols in art design under the background of artificial intelligence," *Mathematical Problems in Engineering*, vol. 2021, Article ID 1258080, 11 pages, 2021.
- [22] D. West, C. Denny, and R. Ruud, "Case study: integrating artificial intelligence metadata within Paramount's digital asset management system[J]," *Journal of Digital Media Management*, vol. 9, no. 3, pp. 198–208, 2021.

- [23] F. D. Weber and R. Schütte, “State-of-the-art and Adoption of Artificial Intelligence in retailing,” *Digital Policy*, vol. 21, 2019.
- [24] M. H. Maras and A. Alexandrou, “Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos,” *International Journal of Evidence and Proof*, vol. 23, no. 3, pp. 255–262, 2019.
- [25] A. A. A. Ahmed and A. Ganapathy, “Creation of automated content with embedded artificial intelligence: a study on learning management system for educational entrepreneurship,” *Academy of Entrepreneurship Journal*, vol. 27, no. 3, pp. 1–10, 2021.

Research Article

Optimization Mold and Algorithm of Risk Control for Power Grid Corporations Based on Collaborative Filtering Technology

Longxing Chen and Ping Han 

The School of Management, Xi'an Jiaotong University, Xi'an 710061, Shaanxi, China

Correspondence should be addressed to Ping Han; hpca@mail.xjtu.edu.cn

Received 9 June 2022; Accepted 22 June 2022; Published 1 August 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Longxing Chen and Ping Han. 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 ever-changing internal and external environmental factors of enterprises, various uncertainties and risks faced by enterprises are increasing, and the feasibility of financial meltdown is increasing. Research on financial meltdown early warning can help enterprises to prevent the occurrence of peril in advance and take resultful measures to ensure the healthy development of enterprises. If a serious financial meltdown leads to the bankruptcy of enterprises, the financial meltdown is not sudden, but a gradual process. The occurrence of financial meltdown is not only a harbinger, but also predictable. Therefore, it is an urgent question to be solved for listed corporations in China that how to mine the message with early warning function from a large amount of financial data generated in the business process of enterprises. The continuous maturity of data mining technique just solves this question. Based on collaborative filtering technique, this paper analyzes the risk control optimization mold and algorithm of power grid corporations, which is of great signification. After research, this algorithm is 30% better than the traditional algorithm, and it is suitable to be proverbially used.

1. Introduction

In the highly competitive market economy, it is not uncommon for listed corporations to get into trouble or even declare bankruptcy due to financial exposures. China's stock market has experienced nearly ten years of development [1]. Faced with the ups and downs of the stock market, some listed corporations have changed from star stocks to junk stocks [2]. Faced with the increasingly complicated economic environment such as increasing competition, tight monetary policy, loss of investor confidence, and fluctuating capital market, whether an enterprise can continue to operate, and how the parties concerned can obtain, message to respond quickly when encountering difficulties are issues of great concern to the enterprise itself, investors, creditors, and even national regulatory agencies and are also a subject worthy of study [3]. The serious financial exposure of the corporation will not only bring huge losses to the production and operation of the enterprise, but also bring huge threats to the relevant stakeholders [4]. It will bring investors' goal

of maintaining and increasing the value of properties and even lose money; The bank cannot recover the loan on time; government security regulatory authorities are facing the pressure of disorderly security market [5]. Most foreign studies on this issue regard the enterprise's filing for bankruptcy according to the bankruptcy law as a sign to determine the enterprise's financial distress, that is, define the financial distress as enterprise bankruptcy [6]. Enterprise bankruptcy refers to a kind of litigation procedure carried out according to bankruptcy law when a corporation is insolvent or fails to reach an agreement with creditors outside the court [7]. Taking enterprise bankruptcy as a sign of entering into financial distress, it is easier to determine the research sample because there is a clear demarcation line between bankrupt enterprises and other nonbankrupt enterprises [8]. In fact, the discovery of financial exposures of listed corporations is always a process, and it takes a long time from the incubation period to the outbreak period. In this long time, we can make predictions through certain techniques. In other words, the financial exposure of listed

corporations is not sudden, but a process of brewing and development. Listed corporations usually start from a normal financial situation, gradually develop into financial difficulties, and finally fall into financial bankruptcy. Therefore, the occurrence of financial exposures of listed corporations is not only a harbinger, but also completely predictable. Establishing a resultful financial meltdown warning system, judging the corporation's operating status according to the market performance and financial message of listed corporations, and getting the signal of financial deterioration as soon as possible will help relevant stakeholders to make scientific decisions and urge relevant parties to take resultful measures in time to reduce risks and losses [9]. Faced with the increasing harmfulness of financial exposures to listed corporations in China, it is a vital question that most listed corporations in China need to solve how to mine message with early warning function from a large amount of financial data generated in the process of business operation [10]. If we study the bankruptcy behavior of enterprises from the financial aspect, it is obviously inappropriate to confine the questions not only affected by financial factors to the financial domain, and the results will not be satisfactory. Therefore, a better starting point for the study of this question should be to predict whether the corporation's financial situation is healthy, not whether it will go bankrupt [11]. It can be said that the occurrence of financial exposures of any listed corporation is a process of gradual deterioration. It is of great signification that we can find financial exposures in time, reduce losses for relevant investors and creditors, and obtain the greatest benefits, and at the same time enable operators to take resultful measures to improve administration when financial exposures sprout. By analyzing and summarizing the causes of financial exposures of listed corporations and their financial characteristics in different periods, it is of great practical signification for operators, investors, and creditors to protect their own interests, as well as for the security regulatory authorities to monitor the stock market, maintain the fairness of the market, and standardize it [12]. We believe that financial meltdown refers to a kind of financial situation that an enterprise may experience in its production and operation. Due to its poor administration, the enterprise loses its solvency, its financial situation deteriorates, and payment peril occurs. If major strategic adjustments such as reorganization and property injection are not taken, the enterprise will face bankruptcy. Although it seems inevitable that some enterprises will die out in the fierce market competition, the death of enterprises is not a one-off event, but a gradual and continuous accumulation process. According to the famous theory of enterprise life cycle, the development of enterprises can be roughly divided into four stages: budding stage, growing stage, mature stage, and declining stage. When an enterprise enters the declining stage, its products are aging, its technique is outdated, its market is saturated, etc., its competitiveness is weakened, and its profitability and solvency are getting worse and worse. If it fails to make major strategic adjustments in time, it will slowly fall into peril and eventually go bankrupt and liquidate. The continuous maturity of data mining technique

and means has just solved this question and opened up a broad space for its financial exposure early warning.

The innovation of this paper lies in the following:

- (1) Collaborative filtering algorithm is introduced, which is the basis of this article, so we should have a cognition of it. Collaborative filtering technique is the most popular technique in recommendation system, which not only has been deeply improved in academic circles, but also has been well applied in the industry.
- (2) This paper introduces the collaborative filtering algorithm of k-means clustering, which is the concrete means we adopt in collaborative filtering algorithm. Collaborative filtering based on k-means clustering is to cluster consumers according to k-means clustering algorithm. Its purpose is to separate the target consumers and their clusters from the whole consumer set, and to alleviate the questions of poor scalability and weak implementation performance caused by too large consumer set.
- (3) The construction of the mold is introduced, so that we can have an overall understanding of the mold. Financial early warning requires that the financial exposures of enterprises can be predicted, so the selected indicators should also be predictive, that is, the possibility of future financial exposures can be predicted by analyzing the historical data formed in the business activities of enterprises, so that the established early warning mold can really predict the financial exposures of enterprises.

This article is divided into five parts:

The first part is the background introduction and the introduction of this article; the second part is the related research of this paper and also mentions this paper. The third part is about the introduction of collaborative filtering. The fourth part is the mold and experimental results, which is the focus of this paper. The fifth part is the conclusion.

2. Related Work

Jiang suggested that the artificial neural network analysis means should be introduced to overcome the shortcomings of the traditional economic early warning mold when establishing the bank loan risk early warning system [13]. Armstrong suggested using neural network theory to establish financial early warning mold [14]. Colak suggested a new mold—ZETA mold—to predict the financial failure of enterprises more accurately. By analyzing 53 bankrupt enterprises and 58 nonbankrupt enterprises from 1969 to 1975, the results show that this mold is obviously superior to Z mold in 1968 [15]. Kuo suggested that financial meltdown be defined as “an enterprise going into legal bankruptcy” [16]. Yang and Han suggested that the mold based on option theory should be applied to financial distress early warning research [17]. The multivariable linear discriminant mold suggested by Shih has strict requirements for early warning variables; that is, the early warning variables are required to

conform to strict joint normal distribution, but the financial ratio of most enterprises in real economic life cannot meet this requirement [18]. Niu et al. suggested that the financial meltdown should be defined as “a serious liquidation question that cannot be solved unless the operation or structure of an economic entity is restructured on a large scale” [19]. Lu et al. suggested using logistic regression to establish financial early warning mold. The research results show that the scale, financial structure (debt ratio), operating performance (return on properties or working capital ratio), and liquidity (current ratio and quick ratio) of an enterprise are highly correlated with the feasibility of financial exposures [20]. Mansi et al. suggested that the financial meltdown be defined as “an economic phenomenon” in which an enterprise is unable to pay its due debts or expenses, including everything from technological failure of fund administration to bankruptcy and everything in between [21]. Galvez et al. suggested that the neural network system should be used to resultfully predict the bankruptcy of the corporation with an accuracy of up to 97% [22].

Early warning analysis of listed corporations’ financial situation not only is the focus of attention of people from all walks of life, but also has vital guiding signification for the future development of listed corporations. Data mining can process the business message of the enterprise, process the data quickly and efficiently, find out the deterioration of the financial targets, and judge the different stages of the financial meltdown that the enterprise is in. Based on collaborative filtering technique, this paper studies the risk control optimization mold and algorithm of power grid corporations, which is of great signification.

3. Collaborative Filtering Technique

3.1. Collaborative Filtering and Recommendation. With the advent of the age of big data and Internet, it has become an inevitable trend for people to enter the age of message overload. The scale of global data has increased from the original level to the current level. Compared with the structured data, which is easy to store in the past, the proportion of unstructured data such as audio, video, pictures, and geographical location message has gradually increased, reaching about 80%. Internet has its own unique characteristics and advantages. First of all, it can realize complete resource sharing; secondly, it is possible to realize efficient and convenient communication through the Internet; third, the development of Internet technique has achieved better fairness. The rapid growth of message on the Internet, on the one hand, makes people’s access to more and more message resources, which brings great convenience to people. On the other hand, people are often confused when faced with a huge amount of message resources, and they have to spend more time and energy searching for helpful message. The phenomenon of “message overload” is becoming more and more serious. Information overload is not only a unique question of the Internet, but also a similar phenomenon in real life, compared with the limited amount of message that the human brain can handle at the same time. That is to say, how to improve the message

retrieval efficiency in the complicated message has become a research hotspot. The appearance of recommendation system just solves the question of data redundancy caused by long data. On the one hand, from the consumer’s point of view, it can avoid the waste of time and energy caused by searching for complicated data, and improve the consumer’s Internet experience; on the other hand, Internet service providers can accurately recommend the content they want to promote and recommend to consumers, thus realizing low-cost and high-return investment.

Whether the consumer message is obtained by explicit tracking or implicit tracking, it can be represented by a consumer-item scoring matrix R . R is a $m \times n$ -order matrix, where m represents the number of consumers and n represents the number of items. For displaying tracking preference message, each element R_{ij} in the matrix represents the i th consumer’s rating value for the j th item, its possible values are $Z = \{v, \dots, V\} \cup \{\bullet\}$, v is the minimum rating value (usually 1), V is the maximum rating value (usually 5 or 7), and \bullet indicates that the consumer has not rated the item as shown in Table 1.

Let the similarity $\text{sim}(u_a, u_i)$ among the consumers in the nearest neighbor sets $U_n = \{u_1, u_2, \dots, u_k\}$, $u_a \notin U_n$, U_n , and u_i ($1 \leq i \leq k$) of the consumer group u_0 with higher similarity be arranged in descending order.

Person correlation coefficient is

$$\text{sim}(u, v) = \frac{\sum_{i \in I_{uv}} (R_{u,i} - \bar{R}_u)(R_{v,i} - \bar{R}_v)}{\sqrt{\sum_{i \in I_{uv}} (R_{u,i} - \bar{R}_u)^2} \sqrt{\sum_{i \in I_{uv}} (R_{v,i} - \bar{R}_v)^2}}, \quad (1)$$

where u and v represent two consumers in the consumer space; $\text{sim}(u, v)$ indicates the similarity between u and v ; I_{uv} indicates the common score item set of u and v , namely, $I_{uv} = \{i \in I | r_{u,i} \neq \bullet \wedge r_{v,i} \neq \bullet\}$; $R_{u,i}$ and $R_{v,i}$, respectively, indicate the ratings of u and v for item i ; and \bar{R}_u and \bar{R}_v indicate the average rating of all items in u and v , respectively.

Let u and v represent the scoring vectors of consumers u and v , respectively, then the similarity between u and v is

$$\text{sim}(u, v) = \cos(u, v) = \frac{\sum_{i \in I_{uv}} R_{u,i} R_{v,i}}{\sqrt{\sum_{i \in I_{uv}} R_{u,i}^2} \sqrt{\sum_{i \in I_{uv}} R_{v,i}^2}} \quad (2)$$

The similarity between consumers u and v is calculated by the following formula:

$$\text{sim}(u, v) = 1 - \frac{\sum_{i \in I_{uv}} (N_{u,i} - N_{v,i})^2}{|I_{uv}|}. \quad (3)$$

Personalized recommendation system is an intelligent recommendation system for big data, which can make corresponding personalized decision support for customers of e-commerce platform. Due to the efficient decision-making mechanism and good recommendation performance, personalized recommendation system has been proverbially used in various platforms. Recommendation system is a resultful solution to the question of message overload. According to the characteristics of consumers, it recommends the targets that meet the needs of consumers and realizes personalized service. The fundamental reason

TABLE 1: User-project rating matrix.

	i_1	...	i_j	...	i_n
u_1	$R_{1,1}$...	$R_{1,j}$...	$R_{1,n}$
...
u_i	$R_{i,1}$...	$R_{i,j}$...	$R_{i,n}$
...
u_m	$R_{m,1}$...	$R_{m,j}$...	$R_{m,n}$

why recommendation system can provide consumers with interesting message is that it relies on consumers' historical behaviors to analyze consumers' needs. Simply put, the essence of recommendation is to link external message through certain ways and means. Among them, collaborative filtering technique is the most popular technique in recommendation system, which has been deeply improved in academic circles and well applied in the industry. The most fundamental idea of synergetic coincidence algorithm is to collaborate among consumers, that is, to make recommendations based on the data of consumers with the same attributes or similarities, which will greatly increase the density of data and the credibility of recommended data. By improving these two aspects, the questions of data sparsity and low credibility are solved, thus improving the accuracy of recommendations.

3.2. Collaborative Filtering Algorithm Based on k-Means Clustering. Collaborative filtering algorithm to resultfully alleviate data sparsity and improve recommendation accuracy belongs to collaborative filtering algorithm based on memory. In order to find the neighbor set similar to the target consumer, we need to traverse all consumers in the system. However, with the continuous expansion of the system scale, new consumers and projects will continue to join, and the shortcomings of the system, such as poor scalability, are becoming more and more serious. To alleviate this situation, many scholars put forward collaborative filtering based on mold. The difference between collaborative filtering based on k-means clustering and general collaborative filtering based on consumers lies in the step of mining neighboring consumers. No matter what kind of mold is based on, the common point of the algorithms is to use historical data to train the mold, and the training of the mold is basically done offline. Compared with the collaborative filtering based on memory, the collaborative filtering algorithm based on mold can resultfully alleviate the shortcomings of the former algorithm. Especially, the advantages of clustering technique, such as strong usability and good scalability for large-scale data sets, can resultfully overcome the above disadvantages. So now we will study the collaborative filtering of k-means clustering. Collaborative filtering based on k-means clustering is to cluster consumers according to k-means clustering algorithm. Its purpose is to separate the target consumers and their clusters from the whole consumer set, and to alleviate the questions of poor scalability and weak implementation performance caused by too large consumer set. Clustering analysis is an exploratory analysis, which is a process of dividing chaotic data into different classes or clusters by certain standards. It is based

on the idea of "birds of a feather flock together." Through certain attributes such as Pearson distance, the similarity of targets within a class and the dissimilarity of targets between classes are as large as possible. Finally, through the clustering results, we can find some connection between data attributes.

The following is a detailed description of the algorithm:

Step 1 : Determine k initial centroids from the set M , and set them as Ss_1, Ss_2, \dots, Ss_N , where centroids can be randomly selected.

Step 2 Calculate the distance d between the non-centroid point and each centroid point. The distance can be measured by Euclidean distance, as shown below:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$
(4)

Step 3 Assign the noncentroid point to the set corresponding to the nearest centroid point and set it to $Sn_1, Sn_2 \dots Sn_N$.

Step 4 Repeat steps 2 and 3.

Step 5 Until the centroid does not change or the iteration stop condition is reached, the final clustering result is given, and the algorithm ends.

Clustering algorithm is proverbially used in many domains, such as recommendation system, deep learning, and intelligent computing. The collaborative filtering algorithm based on k-means clustering first divides consumers into K clusters through consumer message, which directly reduces the data base compared with the general collaborative filtering algorithm based on consumers. At the same time, after mining the neighbor consumers, only the results of the first K nearest neighbors (K is far less than the total number of consumers N) are adopted, which reduces the time complexity of the algorithm to some extent. We can intuitively feel the improvement of operation efficiency in the experiment. The advantage of introducing clustering algorithm into recommendation lies in that the search of target consumers can only be carried out in the cluster containing the target consumers, instead of searching in the whole database. Therefore, the recommendation speed can be improved. This algorithm can reduce the consumer base by clustering consumers, thus improving the efficiency of recommendation to a certain extent. Its simple mold is universal, and it shows good scalability on large-scale data sets.

4. Model Establishment and Experimental Results

At present, domestic and foreign researches are mostly devoted to screening financial ratios or variable

combinations that are both explanatory and stable to build molds. Statistical data mining technique plays a vital role in establishing financial meltdown early warning mold, and the early financial meltdown early warning molds are all based on statistical technique. Discriminant analysis is a statistical analysis means that discriminates the categories of research targets. Discriminant analysis must know the classification of the observed targets and some variable values indicating the characteristics of the observed targets. Discriminant analysis is to screen out variables that can provide more message and establish discriminant function, so that the misjudgment rate of the deduced discriminant function when classifying observation samples is minimal. The structure of data mining system is shown in Figure 1.

From previous empirical studies, it can be seen that most researchers are committed to screening financial ratios or variable combinations that are both explanatory and stable to build molds. Although researchers are trying to prove that the variables they use are superior to other research results, there is no consistent conclusion at present, even a consistent means of screening variables. The design of financial early warning index system should be able to fully reflect the financial status of listed corporations in terms of profitability, growth ability, debt paying ability, property administration ability, and cash flow, so as to analyze the capabilities of enterprises in all aspects, so as to reveal and judge the financial status and future development trend of enterprises as comprehensively as possible. Discriminant financial meltdown early warning mold aims to study the classification of two types of corporations. One is financial meltdown corporation, and the other is financial health corporation. Based on this, a discriminant function is established to discriminate and classify any corporation, borrower, and security issuer. The steps of building a financial early warning system are shown in Figure 2.

In order to expand the business scale or meet the needs of business turnover, almost every listed corporation is in debt, but the debt ratio is different. High debt ratio usually becomes risky business, while low debt ratio may fall into the stereotype of conservative business. As there are many indicators of financial exposure warning, it is impossible to include all financial targets in the study. If too many indicators are selected, the resultfulness and practicability of the mold will be affected. The univariate analysis compares the corporation's vital financial ratios with the standards of the same industry to see if there is a big difference between them, and observes the trend of financial ratios to predict the financial meltdown. Regardless of the level of debt, it is the basic premise for listed corporations to repay their debts when they are due. If a listed corporation fails to repay its due debts, it may be taken over by creditors or judged bankrupt by the court. Therefore, the financial targets related to solvency are often used as vital indicators to examine the financial exposures of listed corporations.

Suppose Y is used to record the financial status of listed corporations, the financial meltdown occurs, and it is recorded as $Y = 1$; if the financial meltdown does not happen, it will be recorded as $Y = 0$. Y depends on another unobservable variable V , while V has a certain functional

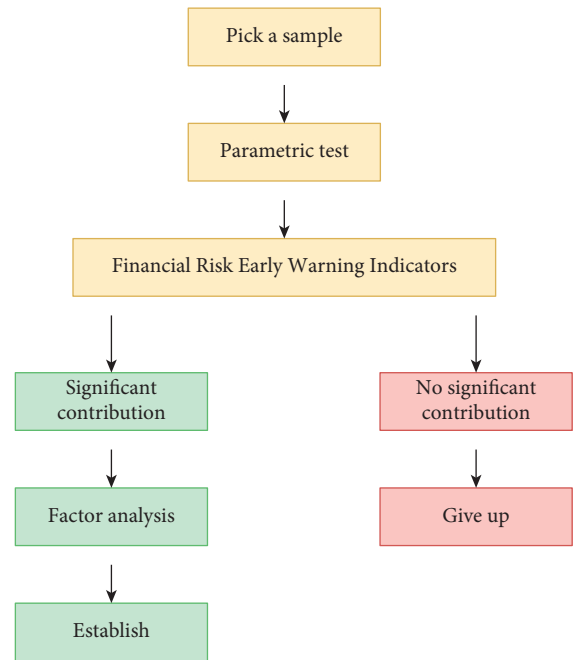


FIGURE 1: Structure of data mining system.

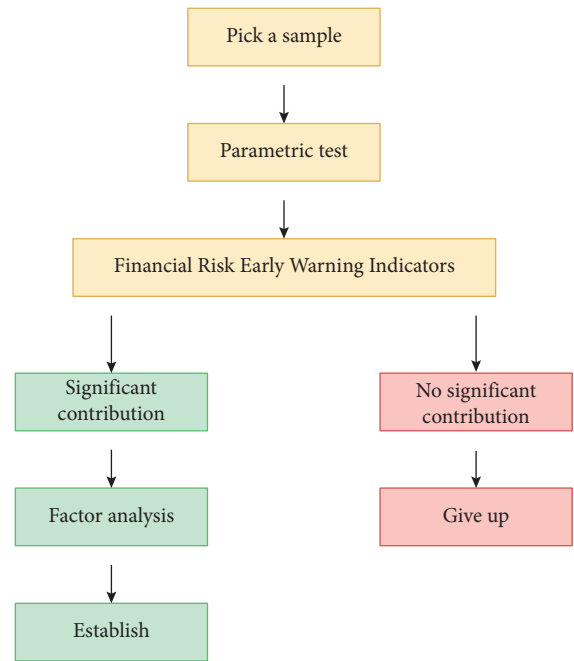


FIGURE 2: Steps to build a financial early warning system.

relationship $V = f(X)$ with the predicted variable X we use, which is assumed to be linear for simple design:

$$v = \beta x_i + \varepsilon_i. \tag{5}$$

The value of V determines whether the event Y occurs, assuming that when $Y = 1$ is equivalent to $V > 0$, then $Y = 0$ is equivalent to $V \leq 0$. Therefore, the feasibility of the event ($Y = 1$) occurring is

$$P(Y = 1) = P(V > 0) = P(\varepsilon_i > -\beta x_i). \quad (6)$$

In order to estimate formula $v = \beta x_i + \varepsilon_i$ according to a certain sample, it is necessary to select a certain feasibility distribution function for ε_i . The feasibility distribution function of ε_i is $F(t)$, and the feasibility distribution function is

$$F(t) = \frac{e^t}{1 + e^t}. \quad (7)$$

From $F(-t) = 1 - F(t)$, formula $P(Y = 1) = P(V > 0) = P(\varepsilon_i > -\beta x_i)$ can be transformed as follows:

$$P(Y = 1) = \frac{e^{\beta x_i}}{1 + e^{\beta x_i}}. \quad (8)$$

When the i th sample is a financial meltdown corporation, $Y_i = 1$. When the i rd sample point is a nonfinancial meltdown corporation, $Y_i = 0$.

$X_i = (X_{i1}, X_{i2}, \dots, X_{ik})$ is the index variable of the i nd sample point:

$$\beta x_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}. \quad (9)$$

From $P(Y = 1) = e^{\beta x_i} / (1 + e^{\beta x_i})$, you can get

$$\ln \frac{P}{1 - P} = \beta x_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}. \quad (10)$$

The selected financial exposure warning indicators should be able to be obtained by calculating the financial data of listed corporations; that is to say, the selected indicators should not only meet the purpose of financial exposure warning, but also be supported by data, and the message required for each data indicator must be easily available. Otherwise, the selected financial meltdown early warning indicators will lose any practical signification because there is no way to calculate them. A failed enterprise has less cash but more accounts receivable. When the cash and accounts receivable are added together and included in liquid properties or current properties, the difference between a failed enterprise and a successful enterprise will be covered up. It is the premise of the existence and development of listed corporations, and the basic condition for listed corporations to survive in the market is to make ends meet. From the market point of view, a long-term loss-making listed corporation always withdraws from the market when it exhausts all its resources. Investors, creditors, managers of listed corporations, and even government managers are increasingly paying attention to the profitability of listed corporations. As shown in Figures 3–5, the mold-based collaborative filtering algorithm used in this algorithm is superior to the traditional content-based collaborative filtering algorithm. With the increase of game times and time, the feasibility of fraud by message providers gradually converges.

Financial early warning requires that the financial exposures of enterprises can be predicted, so the selected indicators should also be predictive; that is, the possibility of future financial exposures can be predicted by analyzing the

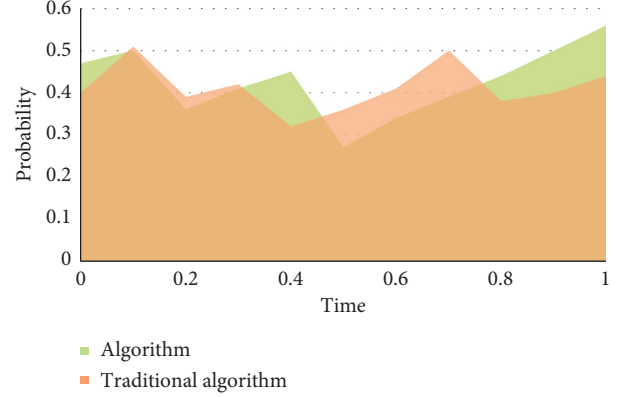


FIGURE 3: Change of equilibrium point.

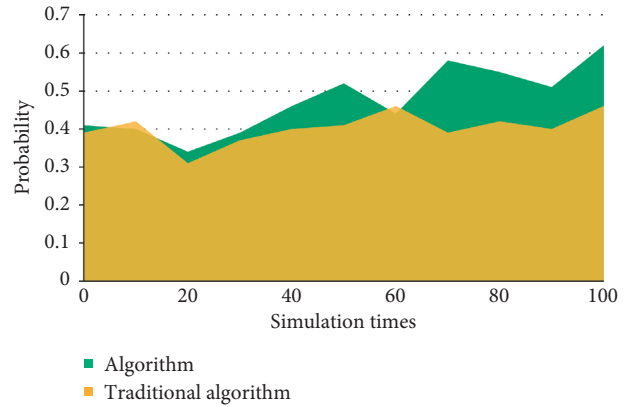


FIGURE 4: Game process.

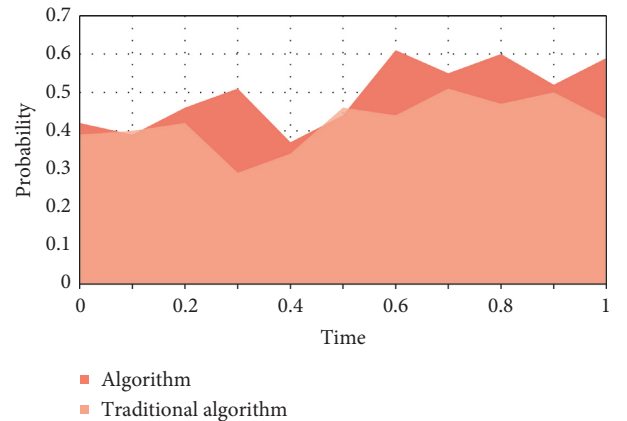


FIGURE 5: Game equilibrium point change.

historical data formed in the business activities of enterprises, so that the established early warning mold can really predict the financial exposures of enterprises. Taking total liabilities as the base is to consider the transformation relationship between long-term liabilities and current liabilities. However, the total liabilities only consider the scale of liabilities, but not the liquidity of liabilities, that is, the debt structure of enterprises. Therefore, there is a great

TABLE 2: Income comparison.

	0	10	20	30	40	50	60	70	80	90	100
Algorithm	56	55	42	39	46	45	51	57	59	60	62
Traditional algorithm	49	50	43	51	53	41	37	52	50	46	55

TABLE 3: Response comparison.

	0	10	20	30	40	50	60	70	80	90	100
Algorithm	66	59	60	41	57	55	62	67	61	70	72
Traditional algorithm	57	55	62	49	44	51	53	44	60	54	56

TABLE 4: Lifting diagram.

	0	10	20	30	40	50	60	70	80	90	100
Algorithm	3.7	2.5	3.1	3.5	1.9	2.6	3.3	3.6	2.7	2.5	3.4
Traditional algorithm	2.9	3	3.4	2.5	2.6	1.8	2.4	3.3	1.9	2.4	3.1

misjudgment for some enterprises that are in peril due to short-term solvency. The index of “total properties” does not take into account the components of properties, and different property items play different roles in the business process, which is not conducive to predicting the profitability of enterprise properties. Asset administration ability is used to measure the efficiency of listed corporations in property administration. The profitability of listed corporations is the result, while the property administration ability is the cause. Listed corporations with good performance should have good property administration ability. Therefore, the financial targets reflecting the property administration ability are also vital indicators used to examine the financial exposures of listed corporations. As can be seen from Tables 2–4, Figures 6–8, the algorithm in this article is 30% better than the traditional algorithm, and it is suitable for being proverbially used. When the consumer moves to the right, it can be kept at a highly stable level.

The selected early warning indicators can sensitively reflect the changes of enterprise operation activities; that is, once the risk factors occur, they can be quickly reflected in the indicator values. Multivariate mold is a multivariate function formula established by using the idea of multivariate mold; that is, a variety of financial targets of an enterprise are weighted and aggregated to generate the total discriminant score of the enterprise to predict the possibility of financial meltdown. Multivariate mold can be regarded as an extension of univariate mold; that is, different financial targets are integrated into one mold. The construction of enterprise financial exposure early warning index system should be scientifically designed according to the causes of enterprise financial exposures, and the index system should be able to make a scientific description of each cause. Profitability is the ability of capital appreciation of enterprises, which is the premise of the existence and development of listed corporations and the comprehensive embodiment of financial structure and operating

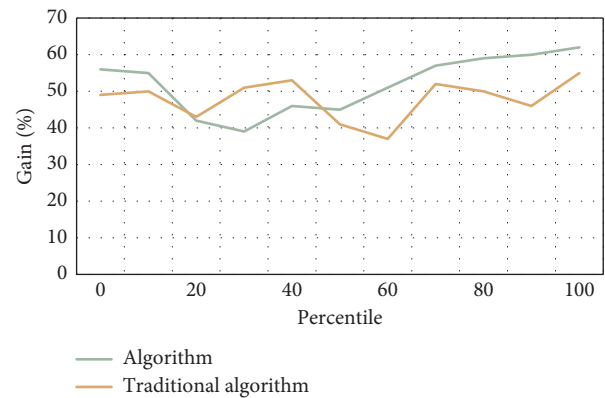


FIGURE 6: Income comparison.

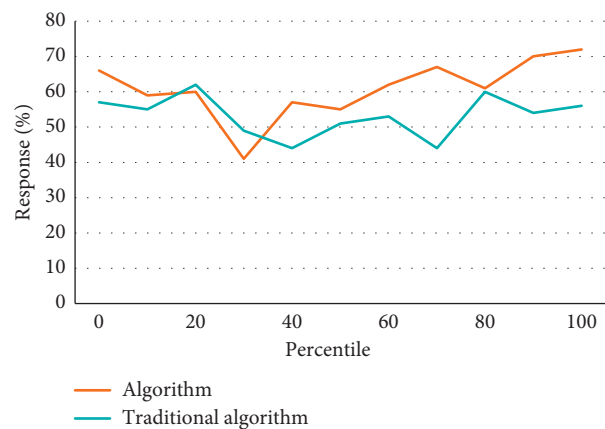


FIGURE 7: Response comparison.

performance. If an enterprise’s profitability is stable, it will have enough surplus to face various possible financial exposures, and the possibility of financial meltdown will be

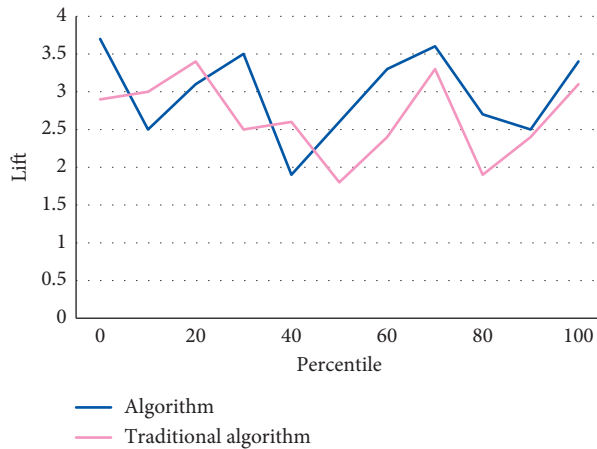


FIGURE 8: Lifting diagram.

lower. Therefore, the financial targets related to the profitability of listed corporations are also used as vital indicators to establish financial exposure early warning molds.

5. Conclusions

The essence of financial distress is the concentrated outbreak of large-scale and high-intensity financial exposures, which is mainly manifested in the extreme deterioration of financial situation, payment peril, and even bankruptcy. Generally speaking, the financial data of listed corporations in China can predict the feasibility of financial exposures in the future; that is to say, the financial targets of listed corporations in China contain certain message content, and consumers of financial data can accurately predict whether financial exposures will occur in the future according to the message provided by the balance sheet, profit statement, and cash flow statement of listed corporations. From the financial point of view, financial distress refers to an enterprise's loss of solvency, that is, its inability to repay its due debts, such as bankruptcy, default on preferred stock dividends, and inability to repay debts. The application of mold-based collaborative filtering technique to financial message processing and financial meltdown early warning has greatly improved the accuracy of financial meltdown early warning, eliminated many external interference factors, and made more hidden message in financial data appear. After research, this algorithm is 30% better than the traditional algorithm, and it is suitable to be proverbially used.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] G. Dionne, J. P. Gueyie, and M. Mnasri, "Dynamic corporate risk management: motivations and real implications[J]," *Journal of Banking & Finance*, vol. 95, no. OCT, pp. 97–111, 2017.

- [2] S. Park, S. Song, and S. Lee, "Corporate social responsibility and systematic risk of restaurant firms: the moderating role of geographical diversification," *Tourism Management*, vol. 59, no. APR, pp. 610–620, 2017.
- [3] H. Zhai, K. C. Chan, Q. Liu, and K. Liu, "The impact of specialized environmental courts on corporate risk-taking: evidence from a natural experiment in China," *Asia-Pacific Journal of Financial Studies*, vol. 49, no. 1, pp. 99–118, 2020.
- [4] B. P. Yost, "Locked-in: the effect of CEOs' capital gains taxes on corporate risk-taking," *The Accounting Review*, vol. 93, no. 5, pp. 325–358, 2018.
- [5] J. P. Haug, U. Pidun, and D. Z. zu Knyphausen-Aufseß, "Cui bono? An empirical investigation into risk benefits of corporate diversification," *Strategic Organization*, vol. 16, no. 4, pp. 429–450, 2018.
- [6] S. P. Ferris, D. Javakhadze, and T. Rajkovic, "An international analysis of CEO social capital and corporate risk-taking," *European Financial Management*, vol. 25, no. 1, pp. 3–37, 2017.
- [7] L. Wei, G. Li, X. Zhu, X. Sun, and J. Li, "Developing a hierarchical system for energy corporate risk factors based on textual risk disclosures," *Energy Economics*, vol. 80, no. MAY, pp. 452–460, 2019.
- [8] L. N. Switzer, Q. Tu, and J. Wang, "Corporate governance and default risk in financial firms over the post-financial crisis period: international evidence," *Journal of International Financial Markets, Institutions and Money*, vol. 52, no. JAN, pp. 196–210, 2018.
- [9] Y. Liu, F. Y. Li, X. Yu, J. Yuan, and D. Zhou, "Assessing the credit risk of corporate bonds based on factor Analysis and logistic regress analysis techniques: evidence from new energy enterprises in China," *Sustainability*, vol. 10, no. 5, p. 1457, 2018.
- [10] J. Poletti-Hughes and G. C. Briano-Turrent, "Gender diversity on the board of directors and corporate risk: a behavioural agency theory perspective," *International Review of Financial Analysis*, vol. 62, no. MAR, pp. 80–90, 2019.
- [11] H. Farag and C. Mallin, "The influence of CEO demographic characteristics on corporate risk-taking: evidence from Chinese IPOs[J]," *The European Journal of Finance*, vol. 24, no. 16-18, pp. 1528–1551, 2018.
- [12] Y. Chu, M. Liu, T. Ma, and X. Li, "Executive compensation and corporate risk-taking: evidence from private loan contracts," *Journal of Corporate Finance*, vol. 64, Article ID 101683, 2020.
- [13] J. Jiang and Y. Chen, "How does labor protection influence corporate risk-taking? Evidence from China," *Pacific-Basin Finance Journal*, vol. 68, no. 1, Article ID 101572, 2021.
- [14] C. S. Armstrong, S. Glaeser, S. Huang, and D. J. Taylor, "The economics of managerial taxes and corporate risk-taking," *The Accounting Review*, vol. 94, no. 1, pp. 1–24, 2019.
- [15] G. Çolak and T. Korkeamäki, "CEO mobility and corporate policy risk," *Journal of Corporate Finance*, vol. 69, no. 4, Article ID 102037, 2021.
- [16] Y. F. Kuo, Y. M. Lin, and H. F. Chien, "Corporate social responsibility, enterprise risk management, and real earnings management: evidence from managerial confidence[J]," *Finance Research Letters*, no. 10, Article ID 101805, 2020.
- [17] O. S. Yang and J.-H. Han, "Exploring the optimal foreign exchange prediction model for corporate foreign exchange risk management: evidence from OECD 10 member states,"

International Business Review, vol. 23, no. 3, pp. 321–350, 2019.

- [18] Y. C. Shih, Y. Wang, R. Zhong, and Y.-M. Ma, “Corporate environmental responsibility and default risk: evidence from China,” *Pacific-Basin Finance Journal*, vol. 68, no. 1, Article ID 101596, 2021.
- [19] G. Niu, L. Yu, G. Z. Fan, and D. Zhang, “Corporate fraud, risk avoidance, and housing investment in China,” *Emerging Markets Review*, vol. 39, no. JUN, pp. 18–33, 2019.
- [20] H. Lu, X. Liu, and L. Falkenberg, “Investigating the impact of corporate social responsibility (CSR) on risk management practices,” *Business & Society*, vol. 61, no. 2, pp. 496–534, 2022.
- [21] S. A. Mansi, Y. Qi, and J. K. Wald, “Bond covenants, bankruptcy risk, and the cost of debt,” *Journal of Corporate Finance*, vol. 66, no. 1, Article ID 101799, 2021.
- [22] J. Galvez, L. Gambacorta, S. Mayordomo, and J. M. Serena, “Dollar borrowing, firm credit risk, and FX-hedged funding opportunities[J],” *Journal of Corporate Finance*, vol. 2021, no. 1, Article ID 101945, 2021.

Research Article

Study of HCP (Hexagonal Close-Packed) Crystal Structure Lattice through Topological Descriptors

Guoping Zhang,^{1,2} Saadia Saeed,³ Adnan Aslam ,⁴ Salma Kanwal ,³ Nazeran Idrees ,⁵ and Tahira Sumbal Shaikh³

¹School of Software, Pingdingshan University, Pingdingshan 467000, China

²Henan International Joint Laboratory for Multidimensional Topology and Carcinogenic Characteristics Analysis of Atmospheric Particulate Matter PM2.5, Pingdingshan 467000, China

³Department of Mathematics, Lahore College for Women University, Lahore, Pakistan

⁴University of Engineering and Technology, Lahore (RCET) 54000, Lahore, Pakistan

⁵Department of Mathematics, Government College University, Faisalabad, Pakistan

Correspondence should be addressed to Salma Kanwal; salma.kanwal@lcwu.edu.pk and Nazeran Idrees; nazeranjawwad@gmail.com

Received 20 April 2022; Accepted 22 June 2022; Published 31 July 2022

Academic Editor: Abdul Rehman Javed

Copyright © 2022 Guoping 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.

Chemical graph theory is a multidisciplinary field where the structure of the molecule is analyzed as a graphical structure. Chemical descriptors are one of the most important ideas employed in chemical graph theory; this is to associate a numerical value with a graph structure that often has correlation with corresponding chemical properties. In this paper, we investigate another very important closed-packed usual crystal structure defined as HCP (Hexagonal Close-Packed) crystal structure and its lattice formed by arranging its unit cells in a dimension for topological descriptors based on a neighborhood degree, reverse degree, and degree. Furthermore, we classify which descriptor is more dominating.

1. Introduction

Chemical graph theory plays a considerable role to investigate a wide range of inorganic and organic chemical structures by studying them through graphical representation. A topological descriptor is a mathematical entity based on certain topological features of chemical structure which correlates with corresponding chemical properties [1–7]. In this research work, we study the most symmetrical and complex close-packed structure defined as HCP(n) using two-dimensional topological descriptors. In graph theory [8], degree of a vertex s in a corresponding graph is the number of edges incident with that vertex defined as d_s . Reverse degree of a vertex s is defined as $\mathbb{R}_s = \Delta(G) + 1 - d_s$, where $\Delta(G)$ is the maximum degree of a vertex in a graph G . Neighborhood degree of a vertex s in G is defined as

$\delta_s = \sum_{t \in N_s} d_t$, where N_s represents the neighborhood degree of a vertex s .

Hexagonal Close Packing (HCP) consists of alternating layers of spheres of atoms (vertices) arranged in a hexagon, with one additional atom (vertex) at the center as shown in Figure 1. Another layer of atoms is sandwiched between these two hexagonal layers which are triangular (three atoms form a triangle by three edges), and the atoms (vertices) of this layer fill the tetrahedral holes created by the top and bottom [9]. The edge set $E(G)$ of HCP contains the edges that connect the atoms that are nearest to each other by an edge. In this context, we represent edges by both a filled and a dotted line to clarify the bonding between two atoms in a 3D HCP(n). The middle layer atoms also share a bond with both hexagonal layers by following a symmetrical pattern.

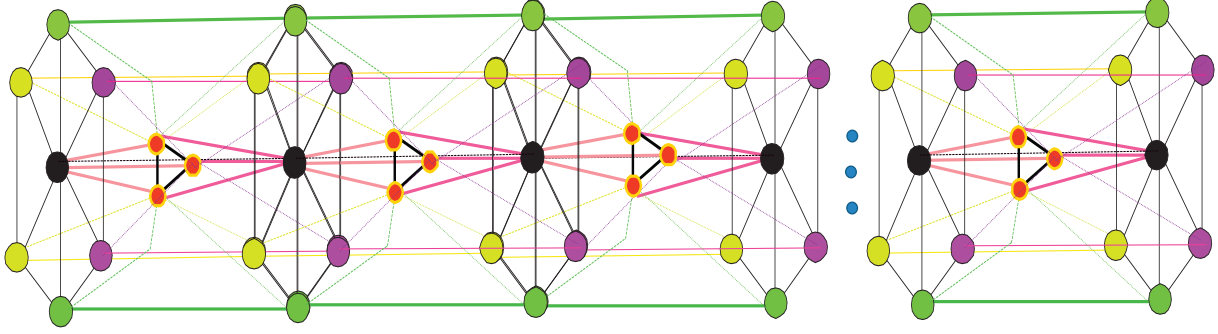


FIGURE 1: HCP (n); hexagonal close packing lattice formed by n unit cells.

Elements that form an HCP crystal structure are zirconium, ruthenium, hafnium, and many more. It is the most strong and brittle structure that is highly found in metals that do not have a very smooth symmetry as in cubic structures, but it contains a vast and stronger metallic property than a usual hexagonal crystal structure.

2. Preliminaries

We give a brief view of some well-known TIs for which we compute closed formulas corresponding to each crystal structure.

Every index is defined as

$$TI(G) = \sum_{\text{edges}} f(H, K). \quad (1)$$

In Table 1, we give a brief view for each descriptor in terms of its function defining in above equation as $f(H, K)$. HCP(n) is a very densely symmetric crystal structure with a complex bonding between its atoms. As it is not a tough thing to study any planar chem structure via descriptors but, here our goal is to perform evaluation for

not a planar but a 3-dimensional structure with a complex symmetry in comparison to cubic structure. We regard the study for the work done to date for complex and 3D crystals, including diamond crystal structures [23], FCC [24], and BCC [25] crystal unit cells that have been investigated by utilizing the definitions from chemical graph theory [5], and allows to study organic chemistry, cordially here, metals and minerals attaining crystal unit cells.

3. Formulation

In this section, we consider that a graph G is HCP(n) defined as hexagonal close-packed crystal structure lattice consisting of n unit cells arranged in one dimension. $V(\text{HCP}(n))$ is defined as vertex set of hexagonal close-packed lattice, and edge set is defined as $E(\text{HCP}(n))$ where for n ordered lattice $V(\text{HCP}(n)) = 10n + 7$, $E(\text{HCP}(n)) = 40n + 12$.

3.1. Degree Based. We utilize Table 2 to compute closed forms for the defined descriptors dependent on degree.

(1) First Zagreb index of HCP(n) is defined as follows:

$$\begin{aligned} M_1(G) &= \sum_{s \in V(G)} (d_s + d_t) = 12(10) + 12(15) + 12(13) + 3n(16) + \\ &\quad \cdot 6(18) + (12n - 18)(14) + (12n - 12)(15) \\ &\quad + (6n - 6)(21) + 2(24) + (n - 2)(28) + \\ &\quad \cdot (6n - 6)(22) + 12(12) = 634n - 10. \end{aligned} \quad (2)$$

(2) Second Zagreb index of HCP(n) is defined as follows:

$$\begin{aligned} M_2(G) &= \sum_{st \in E(G)} d_s d_t = 12(25) + 12(50) + 12(40) + 3n(64) + 6(80) \\ &\quad + (12n - 18)(49) + (12n - 12)(56) + (6n - 6)(98) \\ &\quad + 2(140) + (n - 2)(196) + (6n - 6)(112) \\ &\quad + 12(35) = 2908n - 646. \end{aligned} \quad (3)$$

TABLE 1: Topological indices and their corresponding functions.

Founders	Name	Defining H and K	$TI = f(H, K)$
Gutman [10]	First Zagreb		$M_1(G) = (H + K)$
	Second Zagreb		$M_2(G) = (H \times K)$
Bollobás et al. and Amić et al. [11, 12]	General Randić index	$H = d_s$ and $K = d_t$	$R_\alpha(G) = (H \times K)^\alpha$
Estrada et al. [13]	Atom bond connectivity index		$ABC(G) = \sqrt{H + K - 2/H \times K}$
Vukičević et al. [14]	Geometric arithmetic index		$GA(G) = 2\sqrt{H \times K/H + K}$
Chen et al. [15]	Exponential reduced Zagreb index		$e^{RM2}(G) = e^{(H-1)(K-1)}$
Zhao et al. [16]	General Reverse Randić index, where $\alpha \in \mathbb{R}$	$H = \mathfrak{R}_s$ and $K = \mathfrak{R}_t$	$R_\alpha(G) = [H \times K]^\alpha$
Estrada et al. [13]	Reverse Atom Bond Connectivity index		$\mathbb{R}ABC(G) = \sqrt{H + K - 2/H \times K}$
Vukičević et al. [13]	Reverse geometric arithmetic index		$\mathbb{R}GA(G) = 2\sqrt{H \times K/H + K}$
Shirdel et al. [17]	Reverse hyper Zagreb index		$\mathbb{R}HM(G) = (H + K)^2$
Ghorbani et al. [18]	Neighborhood versio of first Zagreb index		$M_1^*(G) = (H + K)$
	Neighborhood second Zagreb index		$M_2^*(G) = (H \times K)$
Monadal [19]	Neighborhood Forgotten index		$F_N^*(G) = (H^2 + K^2)$
Siddiqui et al. [20]	Neighborhood second modified Zagreb index		$M_2^{mm}(G) = 1/H \times K$
Pal et al. [21]	Third $N De$	$H = \delta_s$ and $K = \delta_t$	$ND_3 = HK(H + K)$
	Fifth $N De$		$ND_5 = [H/K + K/H]$
	Neighborhood Harmonic index		$NH(G) = 2/H + K$
Hosamani et al. [22]	Neighborhood Inverse sum index		$NI(G) = H \times K/H + K$
	The Sanskruti index		$S(G) = (H \times K/H + K - 2)^3$

TABLE 2: Edge partition of $HCP(n)$ based on degree of $V(HCP(n))$

Edge representation	(d_s, d_t)	Cardinality of (d_s, d_t)
E_1	(5, 5)	12
E_2	(5, 10)	12
E_3	(5, 8)	12
E_4	(8, 8)	$3n$
E_5	(8, 10)	6
E_6	(7, 7)	$12n - 18$
E_7	(7, 8)	$12n - 12$
E_8	(7, 14)	$6n - 6$
E_9	(10, 14)	2
E_{10}	(14, 14)	$n - 2$
E_{11}	(8, 14)	$6n - 6$
E_{12}	(5, 7)	12

(3) General Randić index of $HCP(n)$ is defined as follows:

$$\begin{aligned}
 R_\alpha(G) &= \sum_{st \in E(G)} (d_s d_t)^\alpha = 12(d_s d_t)^\alpha \\
 &+ 12(d_s d_t)^\alpha + 12(d_s d_t)^\alpha \\
 &+ 3n(d_s d_t)^\alpha + 6(d_s d_t)^\alpha + \\
 &\cdot (12n - 18)(d_s d_t)^\alpha + (12n - 12)(d_s d_t)^\alpha + \\
 &\cdot (6n - 6)(d_s d_t)^\alpha + 2(d_s d_t)^\alpha + (n - 2)(d_s d_t)^\alpha + \\
 &\cdot (6n - 6)(d_s d_t)^\alpha + 12(d_s d_t)^\alpha.
 \end{aligned}
 \tag{4}$$

$$\begin{aligned}
 R_1(G) &= \sum_{st \in E(G)} (d_s d_t)^1 = 12(25)^1 + 12(50)^1 \\
 &+ 12(40)^1 + 3n(64)^1 + 6(80)^1 \\
 &+ (12n - 18)(49)^1 \\
 &+ (12n - 12)(56)^1 + (6n - 6)(98)^1 + 2(140)^1 \\
 &+ (n - 2)(196)^1 + (6n - 6)(112)^1 + 12(35)^1 \\
 &= M_2(G).
 \end{aligned}
 \tag{5}$$

$$\begin{aligned}
 R_{-1}(G) &= \sum_{st \in E(G)} (d_s d_t)^{-1} = 12(25)^{-1} + \\
 &\cdot 12(50)^{-1} + 12(40)^{-1} + 3n(64)^{-1} + \\
 &\cdot 6(80)^{-1} + (12n - 18)(49)^{-1} \\
 &+ (12n - 12)(56)^{-1} + (6n - 6)(98)^{-1} \\
 &+ 2(140)^{-1} \\
 &+ (n - 2)(196)^{-1} + (6n - 6)(112)^{-1} \\
 &+ 12(35)^{-1} = 0.7455102041 + 0.5790816327n.
 \end{aligned}
 \tag{6}$$

$$\begin{aligned}
 R_{\frac{1}{2}}(G) &= \sum_{st \in E(G)} (d_s d_t)^{\frac{1}{2}} = 12(25)^{\frac{1}{2}} + \\
 &\cdot 12(50)^{\frac{1}{2}} + 12(40)^{\frac{1}{2}} + 3n(64)^{\frac{1}{2}} + 6(80)^{\frac{1}{2}} + \\
 &\cdot (12n - 18)(49)^{\frac{1}{2}} + (12n - 12)(56)^{\frac{1}{2}} \\
 &+ (6n - 6)(98)^{\frac{1}{2}} +
 \end{aligned}$$

$$\begin{aligned}
& \cdot 2(140)\sqrt{\frac{1}{2}} + (n-2)(196)\sqrt{\frac{1}{2}} \\
& + (6n-6)(112)\sqrt{\frac{1}{2}} + 12(35)\sqrt{\frac{1}{2}} = 2.375607208 \\
& + 334.6947784n. \tag{7} \\
R_{\frac{1}{2}}(G) &= \sum_{st \in E(G)} (d_s d_t)^{-\frac{1}{2}} = 12(25)^{-\frac{1}{2}} + 12(50)^{-\frac{1}{2}} \\
& + 12(40)^{-\frac{1}{2}} + 3n(64)^{-\frac{1}{2}} + 6(80)^{-\frac{1}{2}} + \\
& \cdot (12n-18)(49)^{-\frac{1}{2}} + (12n-12)(56)^{-\frac{1}{2}} \\
& + (6n-6)(98)^{-\frac{1}{2}} + \\
& \cdot 2(140)^{-\frac{1}{2}} + (n-2)(196)^{-\frac{1}{2}} \\
& + (6n-6)(112)^{-\frac{1}{2}} + 12(35)^{-\frac{1}{2}} \\
& = 3.371752924 + 4.937319973n. \tag{8}
\end{aligned}$$

(3a) When $\alpha = 1$,

(3b) When $\alpha = -1$,

(3c) When $\alpha = 1/2$,

(3d) When $\alpha = -1/2$,

(4) Atom bond connectivity index of **HCP**(n) is defined as follows:

$$\begin{aligned}
ABC(G) &= \sum_{st \in E(G)} \sqrt{\frac{d_s + d_t - 2}{d_s d_t}} = 12\left(\sqrt{\frac{8}{25}}\right) \\
& + 12\left(\sqrt{\frac{13}{50}}\right) + 12\left(\sqrt{\frac{11}{40}}\right) + 3n\left(\sqrt{\frac{7}{32}}\right) \\
& + 6\left(\sqrt{\frac{1}{5}}\right) + (12n-18)\left(\sqrt{\frac{12}{49}}\right) \\
& + (12n-12)\left(\sqrt{\frac{13}{56}}\right) \\
& + (6n-6)\left(\sqrt{\frac{19}{98}}\right) + 2\left(\sqrt{\frac{11}{70}}\right) \\
& + (n-2)\left(\sqrt{\frac{13}{98}}\right) + (6n-6)\left(\sqrt{\frac{5}{28}}\right) \\
& + 12\left(\sqrt{\frac{2}{7}}\right) = 8.495057575 + 18.66489626n. \tag{9}
\end{aligned}$$

(5) The geometric arithmetic index of **HCP**(n) is defined as follows:

$$\begin{aligned}
GA(G) &= \sum_{st \in E(G)} \frac{2\sqrt{d_s d_t}}{d_s + d_t} = 12\left(\frac{2\sqrt{25}}{10}\right) + 12\left(\frac{2\sqrt{50}}{15}\right) \\
& + 12\left(\frac{2\sqrt{40}}{13}\right) + 3n\left(\frac{2\sqrt{64}}{16}\right) + 6\left(\frac{2\sqrt{80}}{18}\right) \\
& + (12n-18)\left(\frac{2\sqrt{49}}{14}\right) \\
& + (12n-12)\left(\frac{2\sqrt{56}}{15}\right) + (6n-6) \\
& \cdot \left(\frac{2\sqrt{98}}{21}\right) + 2\left(\frac{2\sqrt{140}}{24}\right) \\
& + (n-2)\left(\frac{2\sqrt{196}}{28}\right) + (6n-6)\left(\frac{2\sqrt{112}}{22}\right) \\
& + 12\left(\frac{2\sqrt{35}}{12}\right) = 39.4027062n + 11.35413853. \tag{10}
\end{aligned}$$

(6) Exponential Reduced Zagreb index of **HCP**(n) is defined as follows:

$$\begin{aligned}
e^{RM_2}(G) &= \sum_{st \in E(G)} e^{(d_s-1)(d_t-1)} = 12(e^{16}) + 12(e^{36}) \\
& + 12(e^{28}) + 3n(e^{49}) + 6(e^{63}) \\
& + (12n-18)(e^{36}) + (12n-12) \\
& \cdot (e^{42}) + (6n-6)(e^{78}) \\
& + 2(e^{117}) + (n-2)(e^{169}) \\
& + (6n-6)(e^{91}) + 12(e^{24}) \\
& = 2.487524928 \times 10^{73}n - 4.975049857 \times 10^{73}. \tag{11}
\end{aligned}$$

3.2. *Reverse Degree Based.* We utilize Table 3 to compute closed forms for the defined descriptors dependent on reverse degree of a vertex set of **HCP**(n), where maximum degree for **HCP**(n) = $\Delta(\text{HCP}(n)) = 14$, which we use to compute reverse degree for the entire vertex set of lattice.

(1) Reverse Randić index of **HCP**(n) is as follows where $\alpha = \pm 1, \pm 1/2$:

(1a) When $\alpha = 1$,

TABLE 3: Edge partition of HCP(n) based on reverse degree of $V(\text{HCP}(n))$

Edge representation	$(\mathbb{R}_s, \mathbb{R}_t)$	Cardinality of $(\mathbb{R}_s, \mathbb{R}_t)$
E_1	(10, 10)	12
E_2	(10, 5)	12
E_3	(10, 7)	12
E_4	(7, 7)	$3n$
E_5	(7, 5)	6
E_6	(8, 8)	$12n - 18$
E_7	(8, 7)	$12n - 12$
E_8	(8, 1)	$6n - 6$
E_9	(5, 1)	2
E_{10}	(1, 1)	$n - 2$
E_{11}	(7, 1)	$6n - 6$
E_{12}	(10, 8)	12

$$\begin{aligned}
\mathbb{R}R_1(G) &= \sum_{st \in E(G)} [\mathbb{R}_s \mathbb{R}_t]^1 = \sum_{st \in E_1} [\mathbb{R}_s \mathbb{R}_t] + \sum_{st \in E_2} [\mathbb{R}_s \mathbb{R}_t] + \cdots + \sum_{st \in E_{12}} [\mathbb{R}_s \mathbb{R}_t] \\
&= 12(10 \times 10) + 12(10 \times 5) + 12(10 \times 7) + 3n(7 \times 7) + 6(7 \times 5) \\
&\quad + (12n - 18)(8 \times 8) + (12n - 12)(8 \times 7) + (6n - 6)(8 \times 1) + 2(5 \times 1) \\
&\quad + (n - 2)(1 \times 1) + (6n - 6)(7 \times 1) + 12(10 \times 8) \\
&= 1904 + 1678n.
\end{aligned} \tag{12}$$

(1b) When $\alpha = -1$,

$$\begin{aligned}
\mathbb{R}R_{-1}(G) &= \sum_{st \in E(G)} [\mathbb{R}_s \mathbb{R}_t]^{-1} = \sum_{st \in E_1} [\mathbb{R}_s \mathbb{R}_t]^{-1} + \sum_{st \in E_2} [\mathbb{R}_s \mathbb{R}_t]^{-1} + \cdots + \sum_{st \in E_{12}} [\mathbb{R}_s \mathbb{R}_t]^{-1} \\
&= 12(10 \times 10)^{-1} + 12(10 \times 5)^{-1} + 12(10 \times 7)^{-1} + 3n(7 \times 7)^{-1} \\
&\quad + 6(7 \times 5)^{-1} + (12n - 18)(8 \times 8)^{-1} + (12n - 12)(8 \times 7)^{-1} + (6n - 6)(8 \times 1)^{-1} \\
&\quad + 2(5 \times 1)^{-1} + (n - 2)(1 \times 1)^{-1} + (6n - 6)(7 \times 1)^{-1} + 12(10 \times 8)^{-1} \\
&= \frac{2407n}{784} - \frac{15959}{5600}.
\end{aligned} \tag{13}$$

(1c) When $\alpha = 1/2$,

$$\begin{aligned}
\mathbb{R}R_{\frac{1}{2}}(G) &= \sum_{st \in E(G)} [\mathbb{R}_s \mathbb{R}_t]^{\frac{1}{2}} = \sum_{st \in E_1} [\mathbb{R}_s \mathbb{R}_t]^{\frac{1}{2}} + \sum_{st \in E_2} [\mathbb{R}_s \mathbb{R}_t]^{\frac{1}{2}} + \cdots + \sum_{st \in E_{12}} [\mathbb{R}_s \mathbb{R}_t]^{\frac{1}{2}} \\
&= 12(10 \times 10)^{\frac{1}{2}} + 12(10 \times 5)^{\frac{1}{2}} + 12(10 \times 7)^{\frac{1}{2}} + 3n(7 \times 7)^{\frac{1}{2}} + 6(7 \times 5)^{\frac{1}{2}} \\
&\quad + (12n - 18)(8 \times 8)^{\frac{1}{2}} + (12n - 12)(8 \times 7)^{\frac{1}{2}} + (6n - 6)(8 \times 1)^{\frac{1}{2}} + 2(5 \times 1)^{\frac{1}{2}} \\
&\quad + (n - 2)(1 \times 1)^{\frac{1}{2}} + (6n - 6)(7 \times 1)^{\frac{1}{2}} + 12(10 \times 8)^{\frac{1}{2}} \\
&= 240.6448479n + 183.9070466.
\end{aligned} \tag{14}$$

(1d) When $\alpha = -1/2$,

$$\begin{aligned}
 \mathbb{R}R_{\frac{1}{2}}(G) &= \sum_{st \in E(G)} [\mathbb{R}_s \mathbb{R}_t]^{-\frac{1}{2}} = \sum_{st \in E_1} [\mathbb{R}_s \mathbb{R}_t]^{-\frac{1}{2}} + \sum_{st \in E_2} [\mathbb{R}_s \mathbb{R}_t]^{-\frac{1}{2}} + \dots + \sum_{st \in E_{12}} [\mathbb{R}_s \mathbb{R}_t]^{-\frac{1}{2}} \\
 &= 12(10 \times 10)^{-\frac{1}{2}} + 12(10 \times 5)^{-\frac{1}{2}} + 12(10 \times 7)^{-\frac{1}{2}} + 3n(7 \times 7)^{-\frac{1}{2}} + 6(7 \times 5)^{-\frac{1}{2}} + \\
 &\quad \cdot (12n - 18)(8 \times 8)^{-\frac{1}{2}} + (12n - 12)(8 \times 7)^{-\frac{1}{2}} + (6n - 6)(8 \times 1)^{-\frac{1}{2}} + 2(5 \times 1)^{-\frac{1}{2}} + (n - 2)(1 \times 1)^{-\frac{1}{2}} \\
 &\quad + (6n - 6)(7 \times 1)^{-\frac{1}{2}} + 12(10 \times 8)^{-\frac{1}{2}} \\
 &= 8.921246062n - 3.503847478.
 \end{aligned} \tag{15}$$

(2) Reverse atom bond connectivity index of **HCP** (**n**) is defined as follows:

$$\begin{aligned}
 \mathbb{R}ABC(G) &= \sum_{st \in E_1} \sqrt{\frac{\mathbb{R}_s + \mathbb{R}_t - 2}{\mathbb{R}_s \mathbb{R}_t}} = \sum_{st \in E_2} \sqrt{\frac{\mathbb{R}_s + \mathbb{R}_t - 2}{\mathbb{R}_s \mathbb{R}_t}} + \sum_{st \in E_3} \sqrt{\frac{\mathbb{R}_s + \mathbb{R}_t - 2}{\mathbb{R}_s \mathbb{R}_t}} + \dots + \sum_{st \in E_{12}} \sqrt{\frac{\mathbb{R}_s + \mathbb{R}_t - 2}{\mathbb{R}_s \mathbb{R}_t}} \\
 &= 12 \left(\sqrt{\frac{10 + 10 - 2}{10 \times 10}} \right) + 12 \left(\sqrt{\frac{10 + 5 - 2}{10 \times 5}} \right) + 12 \left(\sqrt{\frac{10 + 7 - 2}{10 \times 7}} \right) \\
 &\quad + (3n) \left(\sqrt{\frac{7 + 7 - 2}{7 \times 7}} \right) + 6 \left(\sqrt{\frac{7 + 5 - 2}{7 \times 5}} \right) + (12n - 18) \left(\sqrt{\frac{8 + 8 - 2}{8 \times 8}} \right) \\
 &\quad + (12n - 12) \left(\sqrt{\frac{8 + 7 - 2}{8 \times 7}} \right) + (6n - 6) \left(\sqrt{\frac{8 + 1 - 2}{8 \times 1}} \right) + 2 \left(\sqrt{\frac{5 + 1 - 2}{5 \times 1}} \right) \\
 &\quad + (n - 2) \left(\sqrt{\frac{1 + 1 - 2}{1 \times 1}} \right) + (6n - 6) \left(\sqrt{\frac{7 + 1 - 2}{7 \times 2}} \right) + 12 \left(\sqrt{\frac{10 + 8 - 2}{10 \times 8}} \right) \\
 &= 24.04625241n + 1.759584801.
 \end{aligned} \tag{16}$$

(3) Reverse geometric arithmetic index of **HCP** (**n**) is defined as follows:

$$\begin{aligned}
 \mathbb{R}GA(G) &= \sum_{st \in E(G)} \frac{2\sqrt{\mathbb{R}_s \mathbb{R}_t}}{\mathbb{R}_s + \mathbb{R}_t} = \sum_{st \in E_1} \frac{2\sqrt{\mathbb{R}_s \mathbb{R}_t}}{\mathbb{R}_s + \mathbb{R}_t} + \sum_{st \in E_2} \frac{2\sqrt{\mathbb{R}_s \mathbb{R}_t}}{\mathbb{R}_s + \mathbb{R}_t} + \sum_{st \in E_3} \frac{2\sqrt{\mathbb{R}_s \mathbb{R}_t}}{\mathbb{R}_s + \mathbb{R}_t} + \dots + \sum_{st \in E_{12}} \frac{2\sqrt{\mathbb{R}_s \mathbb{R}_t}}{\mathbb{R}_s + \mathbb{R}_t} \\
 &= 12 \left(\frac{2\sqrt{10 \times 10}}{10 + 10} \right) + 12 \left(\frac{2\sqrt{10 \times 5}}{10 + 5} \right) + 12 \left(\frac{2\sqrt{10 \times 7}}{10 + 7} \right) + 3n \left(\frac{2\sqrt{7 \times 7}}{7 + 7} \right) + 6 \left(\frac{2\sqrt{7 \times 5}}{7 + 5} \right) \\
 &\quad + (12n - 18) \left(\frac{2\sqrt{8 \times 8}}{8 + 8} \right) + (12n - 12) \left(\frac{2\sqrt{8 \times 7}}{8 + 7} \right) + (6n - 6) \left(\frac{2\sqrt{8 \times 1}}{8 + 1} \right) + 2 \left(\frac{2\sqrt{5 \times 1}}{5 + 1} \right) \\
 &\quad + (n - 2) \left(\frac{2\sqrt{1 \times 1}}{1 + 1} \right) + (6n - 6) \left(\frac{2\sqrt{7 \times 1}}{7 + 1} \right) \\
 &\quad + 12 \left(\frac{2\sqrt{10 \times 8}}{10 + 8} \right) = 35.71316677n + 14.74470034.
 \end{aligned} \tag{17}$$

(4) Reverse hyper Zagreb index of **HCP(n)** is defined as follows:

$$\begin{aligned}
 \mathbb{RHM}(G) &= \sum_{st \in E(G)} (\mathbb{R}_s + \mathbb{R}_t)^2 = \sum_{st \in E_1} (\mathbb{R}_s + \mathbb{R}_t)^2 + \sum_{st \in E_2} (\mathbb{R}_s + \mathbb{R}_t)^2 + \sum_{st \in E_3} (\mathbb{R}_s + \mathbb{R}_t)^2 \\
 &+ \dots + \sum_{st \in E_{12}} (\mathbb{R}_s + \mathbb{R}_t)^2 = 12(10 + 10)^2 + 12(10 + 5)^2 + 12(10 + 7)^2 \\
 &+ 3n(7 + 7)^2 + 6(7 + 5)^2 \\
 &+ (12n - 18)(8 + 8)^2 + (12n - 12)(8 + 7)^2 \\
 &+ (6n - 6)(8 + 1)^2 + 2(5 + 1)^2 \\
 &+ (n - 2)(1 + 1)^2 + (6n - 6)(7 + 1)^2 + 12(10 + 8)^2 \\
 &= 7234n + 7606.
 \end{aligned} \tag{18}$$

3.3. *Neighborhood Degree Based.* We utilize Table 4 to evaluate closed formulas for the neighborhood degree-based descriptors representing topological properties.

(1) Neighborhood version of first Zagreb index of **HCP(n)** is defined as follows:

$$\begin{aligned}
 M_1^*(G) &= \sum_{st \in E(G)} (\delta_s + \delta_t) = \sum_{st \in E_1} (\delta_s + \delta_t) + \sum_{st \in E_2} (\delta_s + \delta_t) + \sum_{st \in E_3} (\delta_s + \delta_t) + \dots \\
 &+ \sum_{st \in E_{21}} (\delta_s + \delta_t) = 12(70) + 12(103) + 12(99) + 12(91) + 2(182) \\
 &+ 12(112) + 12(120) + 6(178) + 12(170) + 12(128) + 6(186) \\
 &+ 2(232) + 6(128) + (3n - 6)(144) + 12(114) + (12n - 42)(116) \\
 &+ (12n - 36)(130) + (6n - 18)(176) + (6n - 18)(190) \\
 &+ 6(132) + (n - 4)(236) = 4792 + 5816n.
 \end{aligned} \tag{19}$$

(2) The neighborhood second Zagreb index of **HCP(n)** is defined as follows:

$$\begin{aligned}
 M_2^*(G) &= \sum_{st \in E(G)} (\delta_s \delta_t) = \sum_{st \in E_1} (\delta_s \delta_t) + \sum_{st \in E_2} (\delta_s \delta_t) + \dots + \\
 &\cdot \sum_{st \in E_{21}} (\delta_s \delta_t) = 12(1225) + 12(2380)12(2240) \\
 &+ 12(1960)2(7752) + 12(3136) + 12(3584) \\
 &+ 6(7269) + 12(6384) + 12(4032) + 6(8208) + 2(13452) + 6(4096) \\
 &+ (3n - 6)(5184) + 12(3248) + (12n - 42)(3364) \\
 &+ (12n - 36)(4276) + (6n - 18)(6844) + (6n - 18)(8496) \\
 &+ 6(4352) + (n - 4)(13924) = 211996n - 130156.
 \end{aligned} \tag{20}$$

TABLE 4: Edge partition of HCP(n) based on neighborhood degree of $V(\text{HCP}(n))$

Edge representation	(N_s, N_t)	Cardinality of (N_s, N_t)
E_1	(35, 35)	12
E_2	(35, 68)	12
E_3	(35, 64)	12
E_4	(35, 56)	12
E_5	(68, 114)	2
E_6	(56, 56)	12
E_7	(56, 64)	12
E_8	(64, 114)	6
E_9	(56, 114)	12
E_{10}	(56, 72)	12
E_{11}	(114, 72)	6
E_{12}	(114, 118)	2
E_{13}	(64, 64)	6
E_{14}	(72, 72)	$3n - 6$
E_{15}	(58, 56)	12
E_{16}	(58, 58)	$12n - 42$
E_{17}	(58, 72)	$12n - 36$
E_{18}	(58, 118)	$6n - 18$
E_{19}	(118, 72)	$6n - 18$
E_{20}	(68, 64)	6
E_{21}	(118, 118)	$n - 4$

(3) The neighborhood forgotten topological index of **HCP**(n) is defined as follows:

$$\begin{aligned}
 F_N^*(G) &= \sum_{st \in E(G)} (\delta_s^2 \delta_t^2) = \sum_{st \in E_1} (\delta_s^2 \delta_t^2) + \sum_{st \in E_2} (\delta_s^2 \delta_t^2) + \sum_{st \in E_3} (\delta_s^2 \delta_t^2) + \dots + \sum_{st \in E_{21}} (\delta_s^2 \delta_t^2) \\
 &= 12(2450) + 12(5849) \\
 &\quad + 12(5321) + 12(4361) + 2(17620) + 12(6272) \\
 &\quad + 12(7232) + 6(17092) + 12(16132) + 12(8320) \\
 &\quad + 6(18180) + 2(26920) \\
 &\quad + 6(8192) + (3n - 6)(10368) + 12(6500) + (12n - 42)(6728) \\
 &\quad + (12n - 36)(8548) + (6n - 18)(17288) + (6n - 18) \\
 &\quad \cdot (19108) + 6(8720) + (n - 4)(27848) = 460640n - 267604.
 \end{aligned}
 \tag{21}$$

(4) The neighborhood second modified Zagreb index of **HCP**(n) is defined as follows:

$$\begin{aligned}
 M_2^{mm}(G) &= \sum_{st \in E(G)} \frac{1}{\delta_s \delta_t} + \sum_{st \in E_2} \frac{1}{\delta_s \delta_t} + \sum_{st \in E_3} \frac{1}{\delta_s \delta_t} + \dots + \sum_{st \in E_{21}} \frac{1}{\delta_s \delta_t} \\
 &= 129 \left(\frac{1}{1225} \right) + 12 \left(\frac{1}{2380} \right) + 12 \left(\frac{1}{2240} \right) \\
 &\quad + 12 \left(\frac{1}{1960} \right) + 2 \left(\frac{1}{7752} \right) + 12 \left(\frac{1}{3136} \right) + 12 \left(\frac{1}{3584} \right)
 \end{aligned}$$

$$\begin{aligned}
& + 6\left(\frac{1}{7296}\right) + 12\left(\frac{1}{6384}\right) + 12\left(\frac{1}{4032}\right) + 6\left(\frac{1}{8208}\right) + 2\left(\frac{1}{13452}\right) + 6\left(\frac{1}{4096}\right) \\
& + (3n - 6)\left(\frac{1}{5184}\right) + 12\left(\frac{1}{3248}\right) + (12n - 42)\left(\frac{1}{3364}\right) \\
& + (12n - 36)\left(\frac{1}{4176}\right) + (6n - 18)\left(\frac{1}{6844}\right) + (6n - 18)\left(\frac{1}{8496}\right) + 6\left(\frac{1}{4352}\right) \\
& + (n - 4)\left(\frac{1}{13924}\right) = 211996n - 130156.
\end{aligned} \tag{22}$$

(5) The third and fifth NDe index of **HCP** (**n**) is defined as follows:

$$\begin{aligned}
ND_3(G) &= \sum_{st \in E(G)} \delta_s \delta_t (\delta_s + \delta_t) = \sum_{st \in E_1} \delta_s \delta_t (\delta_s + \delta_t) + \sum_{st \in E_2} \delta_s \delta_t (\delta_s + \delta_t) \\
&+ \sum_{st \in E_3} \delta_s \delta_t (\delta_s + \delta_t) + \dots + \sum_{st \in E_{21}} \delta_s \delta_t (\delta_s + \delta_t) \\
&= 12(1225)(70) + 12(2380)(103) + 12(2240)(99) + 12(1960)(91) \\
&+ 2(7752)(182) + 12(3136)(112) + 12(3684)(120) \\
&+ 6(7296)(178) + 12(6384)(677) + 12(4032)(128) + 6(8208)(186) \\
&+ 2(13452)(232) + 6(4096)(128) + (3n - 6)(5184)(144) \\
&+ 12(3248)(114) + (12n - 42)(3364)(116) + (12n - 36)(4176)(130) \\
&+ (6n - 18)(6844)(176) + (6n - 18)(8496)(190) \\
&+ 6(4352)(132) + (n - 4)(13924)(236) = 33635504n + 8961688, \\
ND_5(G) &= \sum_{st \in E(G)} \left[\frac{\delta_s}{\delta_t} + \frac{\delta_t}{\delta_s} \right] = \sum_{st \in E_1} \left[\frac{\delta_s}{\delta_t} + \frac{\delta_t}{\delta_s} \right] + \sum_{st \in E_2} \left[\frac{\delta_s}{\delta_t} + \frac{\delta_t}{\delta_s} \right] \\
&+ \sum_{st \in E_3} \left[\frac{\delta_s}{\delta_t} + \frac{\delta_t}{\delta_s} \right] + \dots + \sum_{st \in E_{21}} \left[\frac{\delta_s}{\delta_t} + \frac{\delta_t}{\delta_s} \right] = 12(2) + 12(2.457563025) \\
&+ 12\left(\frac{5321}{2240}\right) + 12\left(\frac{89}{40}\right) + 2\left(\frac{4405}{1938}\right) + 12(2) \\
&+ 12\left(\frac{113}{56}\right) + 6\left(\frac{4273}{1824}\right) + 12\left(\frac{4033}{1596}\right) \\
&+ 12\left(\frac{130}{63}\right) + 6\left(\frac{505}{228}\right) + 2\left(\frac{6730}{3363}\right) + 6(2) \\
&+ (3n - 6)(2) + 12\left(\frac{1625}{812}\right) + (12n - 42)(2) + (12n - 36)\left(\frac{2137}{1044}\right) \\
&+ (6n - 18)\left(\frac{4322}{1711}\right) + (6n - 18)\left(\frac{4777}{2124}\right) + 6\left(\frac{545}{272}\right) \\
&+ n - 42 = 32.28529324 + 85.21361777n.
\end{aligned} \tag{23}$$

(6) The neighborhood Harmonic index of **HCP**(**n**) is defined as follows:

$$\begin{aligned}
 NH(G) &= \sum_{st \in E(G)} \frac{2}{\delta_s + \delta_t} = \sum_{st \in E_1} \frac{2}{\delta_s + \delta_t} + \sum_{st \in E_2} \frac{2}{\delta_s + \delta_t} + \sum_{st \in E_3} \frac{2}{\delta_s + \delta_t} + \cdots + \sum_{st \in E_{21}} \frac{2}{\delta_s + \delta_t} \\
 &+ 12\left(\frac{1}{35}\right) + 12\left(\frac{2}{103}\right) + 12\left(\frac{2}{99}\right) \\
 &+ 12\left(\frac{2}{91}\right) + 2\left(\frac{1}{91}\right) + 12\left(\frac{1}{56}\right) + 12\left(\frac{1}{60}\right) + 6\left(\frac{1}{89}\right) \\
 &+ 12\left(\frac{1}{85}\right) + 12\left(\frac{1}{64}\right) + 6\left(\frac{1}{93}\right) + 2\left(\frac{1}{116}\right) + 6\left(\frac{1}{64}\right) + (3n-6)\left(\frac{1}{72}\right) + 12\left(\frac{1}{57}\right) \\
 &+ (12n-42)\left(\frac{1}{58}\right) + (12n-36)\left(\frac{1}{65}\right) \\
 &+ (6n-18)\left(\frac{1}{88}\right) + (6n-18)\left(\frac{1}{95}\right) + 6\left(\frac{1}{66}\right) \\
 &+ (n-4)\left(\frac{1}{118}\right) = 0.5729928922n + 0.6020913479.
 \end{aligned} \tag{24}$$

(7) The neighborhood inverse sum index of **HCP**(**n**) is defined as follows:

$$\begin{aligned}
 NI(G) &= \sum_{st \in E(G)} \frac{\delta_s \delta_t}{\delta_s + \delta_t} = \sum_{st \in E_1} \frac{\delta_s \delta_t}{\delta_s + \delta_t} + \sum_{st \in E_2} \frac{\delta_s \delta_t}{\delta_s + \delta_t} + \sum_{st \in E_3} \frac{\delta_s \delta_t}{\delta_s + \delta_t} + \cdots \\
 &+ \sum_{st \in E_{21}} \frac{\delta_s \delta_t}{\delta_s + \delta_t} = 12\left(\frac{35}{2}\right) + 12\left(\frac{2380}{103}\right) + 12\left(\frac{2240}{99}\right) \\
 &+ 12\left(\frac{280}{13}\right) + 2\left(\frac{3876}{91}\right) + 12(28) + 12\left(\frac{448}{15}\right) + 6\left(\frac{3648}{89}\right) \\
 &+ 12\left(\frac{3192}{85}\right) + 12\left(\frac{63}{2}\right) + 6\left(\frac{1368}{31}\right) + 2\left(\frac{3363}{58}\right) + 6(32) + (3n-6)(36) \\
 &+ 12\left(\frac{1624}{57}\right) + (12n-42)(29) + (12n-36)\left(\frac{2088}{65}\right) \\
 &+ (6n-18)\left(\frac{1711}{44}\right) + (6n-18)\left(\frac{4248}{95}\right) + 6\left(\frac{1088}{33}\right) \\
 &+ (n-4)(59) = 1402.089842n - 347.4039608.
 \end{aligned} \tag{25}$$

(8) The Sanskruti index of **HCP**(**n**) is defined as follows:

$$\begin{aligned}
 NI(G) &= \sum_{st \in E(G)} \left(\frac{\delta_s \delta_t}{\delta_s + \delta_t - 2}\right)^3 = \sum_{st \in E_1} \left(\frac{\delta_s \delta_t}{\delta_s + \delta_t - 2}\right)^3 + \sum_{st \in E_2} \left(\frac{\delta_s \delta_t}{\delta_s + \delta_t - 2}\right)^3 \\
 &+ \sum_{st \in E_3} \left(\frac{\delta_s \delta_t}{\delta_s + \delta_t - 2}\right)^3 + \cdots + \sum_{st \in E_{21}} \left(\frac{\delta_s \delta_t}{\delta_s + \delta_t - 2}\right)^3
 \end{aligned}$$

TABLE 5: Analyzing descriptors.

Descriptors	General closed formula for hexagonal crystal structure lattice	$n = 1 =$ one unit cell	Check for positivity (+), negativity (-), and domination
M_1 (HCP (n))	$643n - 10$	633	(+)
R_1 (HCP (n)) = M_2 (HCP (n))	$2908n - 646$	2262	(+)
R_{-1} (HCP (n))	$0.579n + 0.7455$	1.3245	(+)
$R_{1/2}$ (HCP (n))	$334.69n + 2.38$	337.07	(+)
$R_{-1/2}$ (HCP (n))	$4.94n + 3.37$	8.31	(+)
ABC (HCP (n))	$18.66n + 8.495$	27.155	(+)
GA (HCP (n))	$39.4027n + 11.354$	50.7567	(+)
e^{RM2} (G)	$2.4875 \times 10^{23} n - 4.975 \times 10^{73}$	$-2.487524929 \times 10^{73}$	(-)
RR_1 (HCP (n))	$1678n + 1904$	3582	(+)
RR_{-1} (HCP (n))	$2407784n - 15959/5600$	0.2203316	(+)
$RR_{1/2}$ (HCP (n))	$240.6448n + 183.907$	424.5518	(+)
$RR_{-1/2}$ (HCP (n))	$8.92n - 3.50$	5.42	(+)
$RRABC$ (HCP (n))	$24.04n + 1.7595$	25.7995	(+)
$RRGA$ (HCP (n))	$35.71n + 14.7447$	50.4547	(+)
$RRHM$ (HCP (n))	$7234n + 7606$	14840	(+)
M_1^* (HCP (n))	$4792 + 5816n$	10608	(+)
M_2^* (HCP (n))	$211996n - 130156$	81840	(+)
M_N^* (HCP (n))	$460640n - 267604$	193036	(+)
M_2^{hm} (HCP (n))	$211996n - 130156$	81840	(+)
ND_3 (HCP (n))	$33635504n + 8961688$	42597192	(+)
ND_5 (HCP (n))	$32.28529324 + 85.21361777n$	117.4989	(+)
NH (HCP (n))	$0.5729928922n + 0.6020913479$	1.17508424	(+)
NI (HCP (n))	$1402.089842n - 347.4039608$	1054.685881	(+)
S (HCP (n))	$2000631.573n - 1847249.828$	153381.745	(+)

(+) and dominates

$$\begin{aligned}
&= 12\left(\frac{35 \times 35}{35 + 35 - 2}\right)^3 + 12\left(\frac{35 \times 68}{35 + 68 - 2}\right)^3 + 12\left(\frac{35 \times 64}{35 + 64 - 2}\right)^3 \\
&\quad + 12\left(\frac{35 \times 56}{35 + 56 - 2}\right)^3 + 2\left(\frac{68 \times 114}{68 + 114 - 2}\right)^3 + 12\left(\frac{56 \times 56}{56 + 56 - 2}\right)^3 \\
&\quad + 12\left(\frac{56 \times 64}{56 + 64 - 2}\right)^3 + 6\left(\frac{64 \times 114}{64 + 114 - 2}\right)^3 + 12\left(\frac{56 \times 114}{56 + 114 - 2}\right)^3 \\
&\quad + 12\left(\frac{56 \times 72}{56 + 72 - 2}\right)^3 + 6\left(\frac{114 \times 72}{114 + 72 - 2}\right)^3 + 2\left(\frac{114 \times 118}{114 + 118 - 2}\right)^3 \\
&\quad + 6\left(\frac{64 \times 64}{64 + 64 - 2}\right)^3 + (3n - 6)\left(\frac{72 \times 72}{72 + 72 - 2}\right)^3 + 12\left(\frac{58 \times 56}{58 + 56 - 2}\right)^3 \\
&\quad + (12n - 42)\left(\frac{58 \times 58}{58 + 58 - 2}\right)^3 + (12n - 36)\left(\frac{58 \times 72}{58 + 72 - 2}\right)^3 \\
&\quad + (6n - 18)\left(\frac{58 \times 118}{58 + 118 - 2}\right)^3 + (6n - 18)\left(\frac{118 \times 72}{118 + 72 - 2}\right)^3 \\
&\quad + 6\left(\frac{68 \times 64}{68 + 64 - 2}\right)^3 (n - 4)\left(\frac{118 \times 118}{118 + 118 - 2}\right)^3 = 2000631.573n - 1847249.828.
\end{aligned} \tag{26}$$

4. Conclusion

From Table 5, we can exactly say that all the descriptors behave positively except exponential reduced Zagreb index and shown in Table 5 by symbol (+) and (-) as the value n increases and the descriptor ND_3 dominates amongst all the topological descriptors which assure us about most dominant topological property for HCP(n), hexagonal closed-packed crystal structure lattice. We compute all the above calculated descriptors for $n = 1$, i.e., for a unit hexagonal close-packed crystal cell.

Data Availability

The data used to support the findings of this study are cited at relevant places within the articles references.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research is partially supported by the China Henan International Joint Laboratory for Multidimensional Topology and Carcinogenic Characteristics Analysis of Atmospheric Particulate Matter PM2.5.

References

- [1] S. Spiro, "The Wiener index of signed graphs," *Applied Mathematics and Computation*, vol. 416, Article ID 126755, 2022.
- [2] C. Chen, M. Liu, X. Chen, and W. Lin, "On general ABC-type index of connected graphs," *Discrete Applied Mathematics*, vol. 315, pp. 27–35, 2022.
- [3] G. Yu, X. Li, and D. He, "Topological indices based on 2-or 3-eccentricity to predict anti-HIV activity," *Applied Mathematics and Computation*, vol. 416, Article ID 126748, 2022.
- [4] E. Estrada and E. Uriarte, "Recent advances on the role of topological indices in drug discovery research," *Current Medicinal Chemistry*, vol. 8, no. 13, pp. 1573–1588, 2001.
- [5] Y. Yang, H. Liu, H. Wang, and H. Fu, "Subtrees of spiro and polyphenyl hexagonal chains," *Applied Mathematics and Computation*, vol. 268, pp. 547–560, 2015.
- [6] Yu Yang, X. Sun, J. Y. Cao, H. Wang, and X. D. Zhang, "The expected subtree number index in random polyphenylene and spiro chains," *Discrete Applied Mathematics*, vol. 285, pp. 483–492, 2020.
- [7] Y. Yang, X. x Li, M. y Jin, L. Li, H. Wang, and X. D. Zhang, "Enumeration of subtrees and BC-subtrees with maximum degree no more than k in trees," *Theoretical Computer Science*, vol. 892, pp. 258–278, 2021.
- [8] B. Bollobás, "Graph theory," *An introductory course*, Springer, Berlin, Germany, 2012.
- [9] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, p. 663, Springer Verlag, Berlin, Germany, 1989.
- [10] I. Gutman, B. Ruscic, N. Trinajstic, and C. F. Wilcox, "Graph theory and molecular orbitals. XII. Acyclic polyenes," *The Journal of Chemical Physics*, vol. 62, no. 9, pp. 3399–3409, 1975.
- [11] B. Bollobás and P. Erdos, "Graphs of extremal weights," *Ars Combinatoria*, vol. 50, pp. 225–233, 1998.
- [12] D. Amić, D. Bešlo, B. Lucić, S. Nikolić, and N. Trinajstić, "The vertex-connectivity index revisited," *Journal of Chemical Information and Computer Sciences*, vol. 38, no. 5, pp. 819–822, 1998.
- [13] E. Estrada, L. Torres, &L. Rodriguez, and I. Gutman, "An atom-bond connectivity index: modelling the enthalpy of

- formation of alkanes," *Indian Journal of Chemistry*, vol. 37, pp. 849–855, 1998.
- [14] D. Vukičević and B. Furtula, "Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges," *Journal of Mathematical Chemistry*, vol. 46, no. 4, pp. 1369–1376, 2009.
- [15] H. Liu, H. Chen, J. B. Liu, and Z. Tang, "Extremal Trees for the Exponential Reduced Second Zagreb index," *Contribution Math*, vol. 3, 2021.
- [16] D. Zhao, Y. M. Chu, M. K. Siddiqui, K. Ali, M. Nasir, and M. T. Younas, M. Cancan, "On Reverse Degree Based Topological Indices of Polycyclic Metal Organic Network," *Polycyclic Aromat. Compd*, vol. 41, pp. 1–18, 2021.
- [17] G. H. Shirdel, H. Rezapour, and A. M. Sayadi, "The hyper-Zagreb index of graph operations," *Iran. J. Math. Chem*, vol. 4, pp. 213–220, 2013.
- [18] M. Ghorbani and M. A. Hosseinzadeh, "A note of Zagreb indices of nanostar dendrimers," *Optoelectron. Adv. Mater.-Rapid Comm*, vol. 4, pp. 1877–1880, 2010.
- [19] S. Mondal and N. De, A. Pal, "On Some New Neighbourhood Degree Based Indices," 2019, <https://arxiv.org/abs/1906.11215>.
- [20] S. Mondal, M. K. Siddiqui, N. De, and A. Pal, "Neighborhood M-polynomial of crystallographic structures," *Biointerface Res. Appl. Chem*, vol. 11, pp. 9372–9381, 2021.
- [21] S. Mondal, A. Dey, N. De, and A. Pal, "QSPR analysis of some novel neighbourhood degree-based topological descriptors," *Complex. Intell. Syst*, vol. 7, no. 2, pp. 977–996, 2021.
- [22] S. M. Hosamani, "Computing sanskruti index of nanostructures," *J. Appl. Math. Comput.*, vol. 54.
- [23] H. Mujahed and B. Nagy, "Exact formula computing the wiener index on rows of unit cells of the diamond cubic grid connected in a row," *PJTAS*, vol. 2, 2019.
- [24] M. K. Siddiqui, M. Imran, and M. Saeed, "Topological properties of face-centred cubic lattice," *Hacettepe Journal of Mathematics and Statistics*, vol. 49, no. 1, pp. 195–207, 2020.
- [25] H. Mujahed and B. Nagy, "Hyper-wiener index on rows of unit cells of the BCC grid," *C. R. Acad*, vol. 71, no. 5, pp. 675–684, 2018.

Research Article

AIOps Architecture in Data Center Site Infrastructure Monitoring

Wei Dong 

IBM, Shenzhen 518000, China

Correspondence should be addressed to Wei Dong; 2016150218@jou.edu.cn

Received 6 June 2022; Accepted 16 June 2022; Published 20 July 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Wei Dong. 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.

AIOps (artificial intelligence for IT operations) has been growing rapidly in recent years. However, it can be seen that the vast majority of AIOps applications are implemented in the IT domain. In contrast, there are few applications in the data center infrastructure domain. Many real-world practices show that a working architecture or algorithm cannot be directly replicated from other domains due to completely different business scenarios and different data characteristics. In this paper, an AIOps architecture for the data center infrastructure monitoring domain is presented. A proven working architecture is given in terms of core modules, such as technical architecture, machine learning algorithms, big data, and business applications, and details are designed in practice. This paper focuses on the technical part and not on each part of the architecture work. In other words, NFRs (nonfunctional requirements), such as performance, availability, manageability, and security, will not be discussed in this paper.

1. Introduction

AIOps and APM (application performance management IT infrastructure management) systems have evolved rapidly in the last 5 years. However, AIOps technology has not made some significant progress in data center infrastructure management. It seems that many reasons have contributed to this problem. In particular, the first two reasons, the relatively small market size and the lack of AIOps capabilities of most DCIM vendors, are the main factors. On the other hand, technology is another challenge. For example, an algorithm may work well in one particular scenario; however, that algorithm may not perform well in another scenario because the data and features are different. Therefore, the design needs to start from the architecture. In this paper, we propose a complete and detailed AIOps architecture which is designed for data center on-site infrastructure monitoring. The mechanism and architecture design are described by the implementation and real-world performance of a number of different machine learning algorithms, using data center room temperature data as an example. The test data clearly demonstrate the effectiveness of this architecture. In addition, this architecture can be a useful reference for those who want to know how to

implement machine learning algorithms in data center site monitoring systems.

2. Research Background

Typically, base station rooms and data center rooms have 3 characteristics that require high concurrency, high throughput, and other nonservice functions to be met when doing the architecture design. These 3 characteristics are as follows.

2.1. Large-Scale Equipment. Taking Shenzhen as an example, in terms of the number of base station rooms, there were 89,000 4G and 5G base stations in 2019. Assuming that there are 10 devices to be monitored, the size of the devices is 890,000 units. Furthermore, if a telecom operator has a 40% market share, the number of devices is expected to have a huge surge, possibly exceeding 356,000. You can see that the value of the number of devices is very large.

2.2. Various Equipment Types and Many Indicators in Different Equipment. It is possible to classify the facilities of data center rooms into the following 4 categories.

Specifically, there are different equipment in different categories. Even the same facility supplier can produce so many different types of products in a particular category. As shown in Figure 1, the components include the following: power monitoring, environmental monitoring, security monitoring, and * network monitoring (optional); the items include the following: power meters, distribution switches, UPS, diesel generators, distribution cabinets, PDUs, DC power supply systems, STS, and ATS battery packs; temperature and humidity, precision air conditioning, water leakage, server room air conditioning, fresh air units, haze, and harmful gases; CCTV, access control system, fire extinguishing system, lightning detection, and antitheft monitoring. Routers, switches, servers, and firewalls [1], as shown in Table 1.

2.3. High Sampling Frequency. In the power supply and environmental monitoring system, the FSU (field supervision unit) is a device that is mainly used for data acquisition and is designed to perform data acquisition [2]. Usually, it collects hundreds of devices every 100 ms or so, and the SC (supervision center) gets this data from the FSU periodically. In most cases, the sampling frequency of the SC is provided as a configuration. Given its value of 1 time/minute, the reported metrics data will be huge because there are many devices that make up many metrics. In addition, those alarm data or other necessary data, such as event data, should be included.

3. Materials and Methods

3.1. Technical Architecture Evolution

3.1.1. All-in-One Architecture. In earlier years, the architecture of data center monitoring systems was a multi-in-one model [3]. Although logically divided into 3 layers, both front-end applications and back-end applications are tightly integrated in this architecture. Usually, the tendency is to use the same framework for all the development work, including the database; for example, all the application development was done with .Net or MFC + SQL Server around the year 2000.

Figure 1 shows the schematic of the all-in-one architecture.

Benefits are as follows:

- (1) Easy to deploy and operate the work.
- (2) In case of a low workload, this architecture can run well and meet business needs in terms of performance.

Disadvantages are as follows:

- (1) In general, scalability is its weakness. It is difficult for this architecture to accommodate fast-growing data, which is difficult for a vertically scaling server to handle.
- (2) Another problem is performance deterioration. As the amount of data grows, eventually, the amount of data will exceed a certain limit, and the performance

of the database will gradually deteriorate. Problems such as unresponsiveness, latency, and even core downtime often occur posing a threat to the business operations.

3.1.2. Big Data Architecture. After 2010, both big data and distributed technologies have become increasingly popular [4]. In particular, the decoupling of front-end applications and back-end applications is accompanied by the introduction of big data, caching, and other technologies in the architecture. The big data architecture is shown in Figure 2.

Benefits are as follows:

- (1) Spark, HDFS, Kafka, Redis, and MongoDB middleware are widely used in current technology solutions, enhancing throughput and alleviating performance bottlenecks
- (2) Thanks to the natural advantages of distributed and big data technologies, the response time of the reporting module is also greatly improved, especially when querying historical data [5]

Disadvantages are as follows:

It is not efficient enough compared to the progress of APM on AIOps. It lacks artificial intelligence techniques to automate some operational tasks, especially in anomaly detection and root cause analysis work

For example.

- (1) A large amount of configuration work is handled manually.
- (2) The event is usually not noticed until it occurs.

Although CEP (complex event processing) can raise some alerts in advance. Overall, the MTTD (mean time to detection) is still slow. As devices and applications are added, new problems are added, creating serious challenges for modules that have fixed rules centered on user experience.

- (3) Long MTTR (mean time to repair).

RCA (root cause analysis) depends entirely on the skills and experience of engineers, and its time cost varies from person to person. In general, the speed of problem-solving remains low, as shown in Figure 3.

3.1.3. Artificial Intelligence + Big Data. The introduction of artificial intelligence technology can improve the quality of the product and increase productivity [6]. An architecture combined with AI can be a good solution for redundant work in certain scenarios and can also help in reducing MTTD and MTTR. The following is an AIOps architecture which consists of big data, AI, and microservices, as shown in Figure 2.

3.2. Architecture Design

- (1) Mainly, the architecture is for the product. In other words, the architecture is the technical design of the product, including functional and nonfunctional

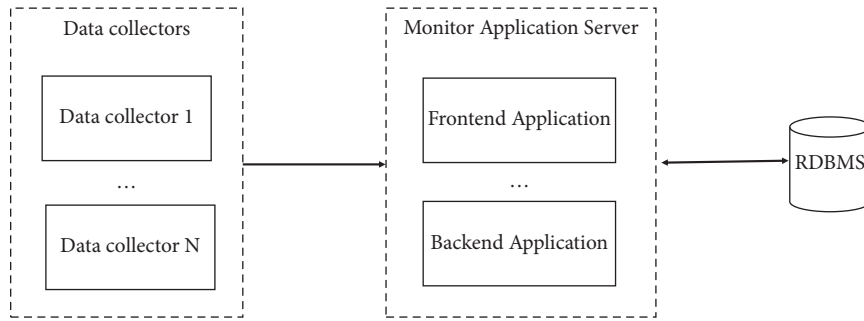


FIGURE 1: Architecture diagram.

TABLE 1: Power and environmental equipment and indicators.

Components	Items
Power monitoring	Coulometer, distribution switch, UPS, diesel generator, power distribution cabinet, PDU, DC power supply system, battery pack, STS, and ATS
Environment monitoring	Temperature and humidity, precision air conditioning, water leakage, computer room air conditioning, fresh air unit, smog, and harmful gas
Security monitoring	CCTV, access control system, fire extinguisher system, lightening protection detection, and antitheft monitoring
*Network monitoring (Optional)	Router, switch, server, and firewall

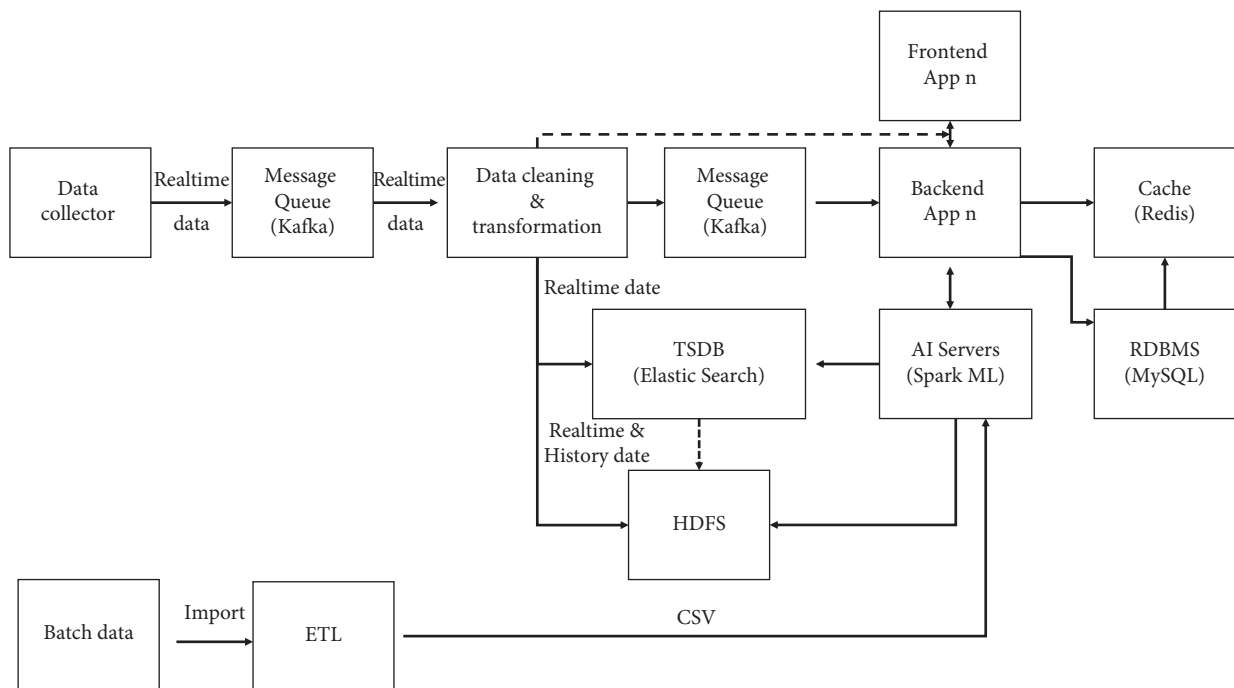


FIGURE 2: Architecture diagram.

features [7]. In general, an AIOps product needs to satisfy these requirements. First, it can be presented to run and delivered as a holistic solution. Second, it can be embedded directly into a web page or other 3rd application. Finally, it is also capable of being served entirely as a back-end application. The advantages of a product with such features are self-

evident. Regardless of what the product is, as a specific solution or being integrated into other products, such a feature would be well suited to meet the flexibility requirements of the market.

- (2) To address the problem of heterogeneous data access by third parties, a mapping table was designed in the database design work [8]. Since the reported data and

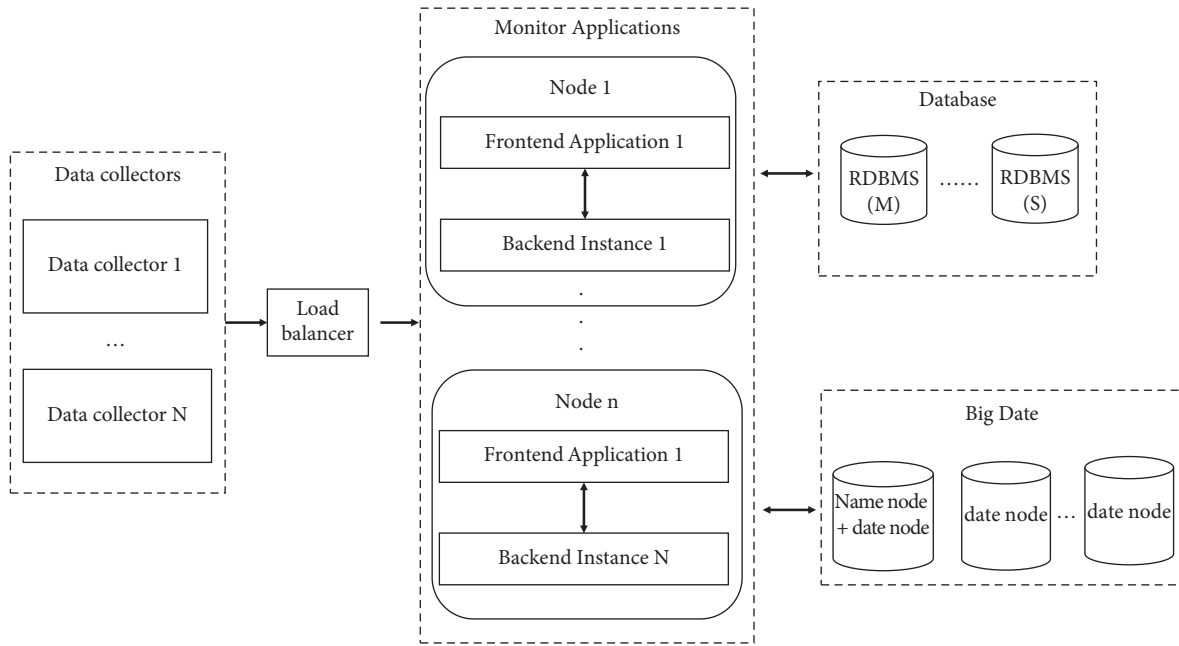


FIGURE 3: Big data architecture.

the internal AIOps data may differ in terms of device types and metrics. A device mapping table is needed with a schema consisting of raw device type, raw metrics, internal device type, and metrics.

As shown in Tables 2 and 3.

- (3) In the data storage section, different data are considered to have different usage frequencies, and these data are classified according to the access frequency. Specifically, hot data, warm data, and cold data are defined as having been accessed at least once in the last 24 hours, 1 month, and more than 1 month, respectively. Moreover, this feature is configurable. On the other hand, different data have different storage media. Accordingly, hot data are stored in Redis and Elasticsearch, warm data is stored in Elasticsearch, and cold data are stored in HDFS in the order given. In contrast, business data are stored in an RDBMS, such as MySQL.

Focusing on the feature that the writing frequency of the performance data is much higher than the reading frequency, the time series database Elasticsearch is adopted to store the performance data and alarm data.

- (4) Processing and storing training data, prediction results, and real-time performance data in different systems. The reason is that we wish to avoid the performance of the system from being affected. Specifically, assuming that the collection, processing, and storage of real-time data are deployed on a single device, this device will become a single point of failure, as this data may be used to support concurrent requests from the business and processed by multiple back-end services, e.g., AI computing modules and back-end applications. Given the high

system performance requirements of the AI module and back-end applications, two different Kafka themes are used to distribute real-time data to Elasticsearch and the HDFS, respectively. For example, both real-time data storage requirements and processing requirements need to be met. And generally, due to decoupling, these two requirements are implemented in two different systems. Therefore, if these two systems are designed to access the same storage medium at the same time, it will create a bottleneck in the system, and the situation will even deteriorate rapidly when dealing with large amounts of real-time data.

- (5) To improve data throughput, Kafka + Spark streams were introduced. Compared with traditional MQ middleware, Kafka has higher performance in terms of reading/writing speed, throughput, and reliability. On the other hand, the Spark streams are definitely better than traditional ETL tools in real-time data processing.
- (6) In the machine learning section, the following features are introduced with Python + scikit-learn.
- (i) Semi-automatic data annotation methods based on machine learning algorithms and manual corrections
 - (ii) Data import. For example, importing data through web pages or web services
 - (iii) Feature selection by expert experience
 - (iv) Algorithm management
 - (v) Automatic optimization of hyperparameters
 - (vi) Model rolling optimization

In addition, for comparison, a deep learning algorithm LSTM is adopted for validation, and the Keras framework is used to implement LSTM.

- (7) In terms of business service applications, business-function-oriented web services are proposed to combine artificial intelligence with business services. Also, these web services are provided on the web accordingly, such as real-time performance data prediction, alarm prediction, AI algorithms, feature management, and training data management. From a technical point of view, Spring Cloud + VUE + Nginx is used to decouple the back-end and front-end. The business services are based on Spring Boot and Spring Cloud microservices.
- (8) For bulk data, an additional ETL toolset is provided to easily import data into HDFS or ES.

3.3. Main Service Flows

- (1) The sensor data collection application sends real-time data to the Kafka message queue.
- (2) The data cleaning and transformation module receives the message, subscribes to the relevant topics, and completes the data cleaning work, e.g., data anomaly detection and missing data processing [9]. In particular, the anomaly detection module will mark the type of data, whether it is normal or not. If the data is anomalous, then both the original data and the corrected data will be saved separately. Usually, these data are saved in Elasticsearch and the HDFS. Also, the data is forwarded to Kafka so that the back-end service can be notified immediately.
- (3) The back-end service gets the real-time data on the subject of Kafka and displays them through a web page. A line graph is proposed in which two lines are drawn, namely, the real-time data line and the predictive data line.

Thus, users can easily switch between different algorithm graphs so that the most suitable algorithm stands out naturally. (b) In addition, the back-end service will analyze the gap between real-time and predicted data and report alerts immediately or not, depending on the results of the analysis of dynamic thresholds.

- (4) To prevent the model drift problem, the hyper-parameters of the AI model will be updated periodically based on the latest data [10]. Therefore, a configuration page is provided to set the frequency of cron jobs, specifying the time, such as hours, days, weeks, or months. For example, a job can be executed every 24 hours. The job will obtain the latest $7 * 24$ data to train the model. Given a device-specific metric, it will calculate and give predicted values for the next 24 hours. These values will be stored in the cache and RDBMS instead of the HDFS, which ensures the timeliness of the system interactions.
- (5) In addition to this, the commercial application, a java back-end service, also provides management and configuration modules such as device management,

metrics management, algorithm management, and model management [11].

4. Results and Discussion

4.1. Technology Selection and Core Implementation

4.1.1. Message Queue Kafka. Compared to other MQs, Kafka is able to handle larger volumes of data and can easily scale without data loss [12]. On the other hand, as a product of the big data ecosystem, Kafka is widely used in various projects. Therefore, it has proven to be stable. Last but not least, it is possible to easily decouple these different services using Kafka, thus enhancing the robustness of the system.

4.1.2. Data Transformation Module. The main functions of the data transformation module are as follows:

- (1) Exception processing of historical data.

Usually, it is not necessary to handle exceptions to historical data immediately [13]. Various machine learning algorithms, such as IForest, KNN, DBScan, and the traditional fixed threshold, are used for exception detection. When the results of these algorithms are incorrect, one can also correct the results of the data labels.

It is also necessary to classify these data by anomaly type [14]. In general, these types of anomalous data are derived and different strategies are adopted to deal with them, as shown in Table 4.

- (2) Exception processing of real-time data. Considering the processing delay, relevant APIs are directly used to get fast results instead of using machine learning algorithms such as IForest, KNN, and DBSCAN [16].

Null and missing data are discarded and stored in the HDFS and Elasticsearch. As a compensating measure, prediction algorithms such as ARIMA, Holt-Winters, and LSTM are used to generate relevant new data to replace the original data.

Contextual anomalous data may be compared to normal data, such as "1," "2," "3"! @#..., marks contextual exception data as an abnormal value. In addition, additional modifications are suggested, for example, replacing it with the latest data.

After the data processing work is finished, the original data and the processed data are stored in the database. This work can then be repeated in the same way when the model is updated with the latest new data [17].

- (3) Conversion of data types

Depending on the system requirements, the raw data need to be converted into internal data for AIOps, as mentioned in the data mapping section above. The raw data will be sent to Kafka. Then, this data will be processed by the ETL application [18]. As a result, the processed data will be distributed among ES, HDFS, and MySQL. Also, the processed data will be

TABLE 2: t_devicetypemap.

Name		t_devicetypemap		
Primary key		id		
Foreign key		oriDevType		
Column	Specification	Type	Default value	Remarks
id	Auto-increment 1	bigint unsigned	Nonempty	
oriDevType	Source device type	varchar (64)	Not null	
sDevType	Destination device type	int	Not null	

TABLE 3: t_metetypemap.

Name		t_metetypemap		
Primary key		id		
Foreign key		oriDevType, oriMtType		
Column	Specification	Type	Default value	Remarks
id	Auto-incremented by 1	bigint unsigned	Not null	
oriDevType	Source device type	varchar (64)	Not null	
Metric type (oriMtType)	Source metric type	varchar (64)	Not null	
sDevType	Decomposition device type	int	Not null	
sMtType	Type of decomposition metric	int	Not null	

sent to Kafka to notify the back-end services to complete the real-time processing on time.

(4) Technical Framework

The ETL application is implemented using Spark streaming and Scala, and the ETL application is treated as a standalone application. The reason why Scala and Kafka are used instead of other languages to interact with ES and MySQL is because Scala has a clear advantage in real-time processing (although Storm is better than Scala in terms of real-time processing in the internal tests conducted, the difference between Storm and Scala is small. Most importantly, Scala’s ecosystem is better than Storm’s ecosystem. In other words, it is a trade-off). In short, everything that is done is to increase the processing speed and reduce the complexity of the technology stack.

4.2. Data Storage. Data storage design is not an easy task. One needs to take into account the business realities and the goals of the project and try to give the best solution.

In this monitoring system, a TSDB is used to store performance data and alarm data. Besides, considering the response latency requirement and the different frequencies between read and write operations, a multitiered storage system was built, which includes TSDB, cache, RDBMS, and HDFS. First, there is the fact that the write time is much more than the read time. TSDB, such as InfluxDB or Elasticsearch, is a better choice compared to those traditional RDBMSs. On the other hand, MySQL is chosen as the RDBMS to store business data because it has excelled in traditional business data storage for the past 20 years. For data that are used frequently, it is stored not only in the RDBMS but also in a cache, such as Redis. For archived data

and training data, HDFS is chosen because it is an inexpensive and reliable solution for storing big data.

4.2.1. TSDB. Choosing a TSDB to store performance data is a natural choice. However, choosing which TSDB product to choose is not an easy task, as one needs to know exactly what one wants to have, and also needs to give comparative dimensions for these different TSDB products.

Mainly, investments were made in InfluxDB, ES, and OpenTSDB.

Although InfluxDB has better performance in either the standalone mode or distributed mode. However, since the distributed version of Influx is a paid software, the switch was quickly made to ES, which is a free distributed-oriented product with fairly high throughput and user-friendly queries. In addition, OpenTSDB was reviewed. However, as known, a few projects did choose OpenTSDB. Unfortunately, many of them abandoned OpenTSDB in favor of ES for various reasons, including some top companies.

4.2.2. Storage Classification and Hierarchy. For business data, the following different business types are present: time-series data (performance data), alarm data, and business data.

These data have different frequencies of use. Therefore, it is necessary to make the frequency configurable to meet different requirements.

Hot data, stored in Redis and ES, are defined as having been accessed at least once in the last 1 week.

Hot data, stored in ES, are defined as having been accessed at least once in the last 1 month. Warm data consists mainly of time-series data, such as performance data and alarm data.

TABLE 4: Handling of exception data.

Type of anomaly	Policy
Point anomalies ([15], “anomaly detection. a survey”)	Marked as an anomaly
Contextual anomalies ([15], “anomaly detection: a survey”)	Discarded directly and replaced with the most recent data.
Null values	Discarded directly and replaced with the most recent data.
Missing values	Discarded directly and replaced with the closest data.
Duplicate values	When the timestamp, object, and location are identical, they are deleted.

Cold data, stored in HDFS, are designed to store archived data or training data. It would be a better choice for storing data that has not been accessed for more than 1 month.

Finally, the business data are treated as a single type and MySQL is used to store it.

4.2.3. Repeated Storage to Avoid SPOF (Single Point of Failure). Assume that both data processing and data storage are installed on one node (server). Since a large amount of data may be reported at the same time, this node may become the SPOF of the system. The reasons are as follows: first, the data may be accessed by those front-end services or users at the same time, but it is also possible that the data are being processed by other services, for example, artificial intelligence applications, which is a time-consuming and high-throughput data access. Therefore, Kafka is used to forward the data to the ES and HDFS (or other storage media) separately.

4.3. Business Applications. Business applications are designed to accomplish business logic. Specifically, it provides web UI, system configuration, machine learning algorithm management, and data management [19].

- (1) Three services are proposed, namely, the java application for AIOps, whose content is the implementation of business logic, the ETL application using Spark Stream technology, implemented by Scala, and the AI application providing machine learning capabilities.
- (2) The spring cloud components are introduced in the java application, such as spring boot, spring gateway, and eureka. Front-end services are written by VUE and deployed on the Nginx server. The front end and java application are decoupled through rest services of the java application.
- (3) These functions are provided in the java application, which consists of three parts. The first part includes service management, configuration, and web services for integration, including real-time data access and batch data access, alarm data processing, and events. The second part is the management of artificial intelligence models, data management, algorithm configuration, and training of models. The third part is the storage design. The database tables are carefully designed

so that data from third parties can be easily integrated into the system.

4.4. Machine Learning Module. Machine Learning Algorithms.

The AI service provides implementations of various machine learning algorithms or deep learning algorithms for anomaly detection, classification, and prediction. It also gives the AIOps system the ability to perform the following functions: semi-automatic annotation, feature selection, data normalization, automatic generation of hyperparameters and their optimization, rolling optimization of model hyperparameters, model evaluation, model deployment, and algorithm management.

4.5. Design Ideas. The following automatic functions need to be provided. In addition to this, a human intervention interface needs to be provided so that one can review or correct the computational results of the model. Otherwise, the diligence requirements will not be met [20].

- (1) Automaticity of the model can be reflected in automatic optimization of hyperparameters, automatic updating of the model, and automatic detection and correction of abnormal data.
- (2) Data competition needs to be considered to avoid AI applications and other applications accessing the same storage medium at the same time, thus mitigating the risk of SPOF.
- (3) In terms of the response time, the system should give a real-time response to the user. For example, from a technical point of view, the predicted values for the next few hours should be given. In this way, the difference between the actual data and the predicted data can be compared immediately when the actual data are available, rather than performing time-consuming prediction calculations in real-time.
- (4) Many machine learning algorithms, especially time series algorithms, require high-quality data. For time-series data, the anomalous data must usually be properly processed and the time span between the anomalous data must be strictly determined before computation. In some cases, some auxiliary processes are even required. For example, often two differentiations are done in ARIMA to obtain a stable data set.

- (5) These algorithms must be configurable so that they can be easily maintained, such as CRUD operations and hyperparameter settings.
- (6) The deployment management of the model should be included; not only the upgrade of the model, but also the fallback when the model fails.
- (7) Selection rules for training data, a mechanism for generating data labels, and archiving historical data should all be established. For example, the training data should preferably be balanced rather than biased, and the batch size should also be set to some extent according to the training speed and accuracy requirements of the model.
- (8) There is no point in telling an algorithm whether it is good or bad without considering the business scenario and the inherent characteristics of the data. In other cases, the most popular algorithm may not be the right algorithm to solve your problem. Instead, you should let the results of the algorithm tell which one is the best. In fact, in many cases, even a simple algorithm may beat a popular algorithm in a given training data, depending on how well the algorithm matches the inherent characteristics of the data.

In order to implement dynamic threshold features for temperature data and to filter the best algorithms, three algorithms, ARIMA, LSTM, and Holt-Winters, are introduced and their performance on the same dataset is compared.

The implementation of ARIMA will be used as an example to show how machine learning algorithms work in a real AIOps project.

4.6. Core Implementation of ARIMA. ARIMA is a traditional time-series algorithm that is widely used for forecasting [21]. Considering the fact that the whole process of ARIMA is an important reference for other machine learning algorithms applied to other projects, some main details of ARIMA implementation will be given.

First, from the perspective of the data center room, the temperature of the room is relatively stable within a certain narrow range compared to the wide variation of temperature in different seasons. On the other hand, many data center specifications have their own standards for the temperature of data center rooms, some of which are mandatory.

Three months of temperature data were taken and analyzed to plot the time series. At that time, there was really no sign of seasonality or trend seen. It is clear from the line graph that there was a rapid and sharp drop in data at the beginning of August, and significant fluctuations in data throughout July, which is reasonable because the data center was in the trial installation phase before August.

Finally, after the trial run phase, the data in August were more stable.

The temperature data changes are shown in Figures 4–7.

However, it was found that although the stable part of these data, such as the data for August, was used, the ADF test was still unsatisfactory. In particular, the p value of ADF

was 0.05 and the ADF value of the data was 0.063154, which is greater than 0.05 and is generally not accepted.

The results are as follows:

```
test statistic -2.766938
the p value is 0.063154
# of lags used 21.000000
number of observations used 816.000000
critical value (1%) -3.438389
critical value (5%) -2.865088
critical value (10%) -2.568660
```

Therefore, single-difference and double-difference experiments were done separately to verify which result would be better. As a result, the results of double-difference experiments were better than those of single-difference experiments. In addition, triple differencing was also tried. Unfortunately, it was found that there was only a very small boost in ADF, and there was a significant processing delay. Therefore, the recommendation is $d = 2$.

On the other hand, there are other hyperparameters, such as p and q . One way to obtain the values of these two parameters is to manually use ACF and PACF. Obviously, such a method is not efficient compared to using the auto-ARIMA function to automatically generate the values of p , d , and q . However, in the project, it was found that if the value of d was set manually by observing the real behavior of the data line. For example, $d = 2$ was initially set after having a clear understanding of the data. Given the value of d , auto-ARIMA can be used to automatically generate p and q and finally get the best combination of p , d , and q .

Moreover, time series algorithms, such as ARIMA, Holt-Winters, and LSTM, need to classify the training data within a fixed time interval. However, in reality, this is almost an impossible task for various reasons.

The answer is to take the last data as the starting data. Then, re-sort the data in the given time interval. For example,

```
Initial data.
2018/8/19 0:49 23
2018/8/19 1:53 23.2
2018/8/19 2:57 23.6
2018/8/20 3:01 23.6
2018/8/19 4:01 24
2018/8/19 5:05 24.4
2018/8/19 6:09 23.1
2018/8/19 7:13 23
2018/8/19 8:17 23.4
2018/8/19 9:21 22.9
2018/8/19 10:25 22.9
2018/8/19 11:29 23.1
2018/8/19 12:33 23.6
```

In the abovementioned data, the last data, “2018/8/19 12:33 23.6,” is used as the starting data, then, a fixed time

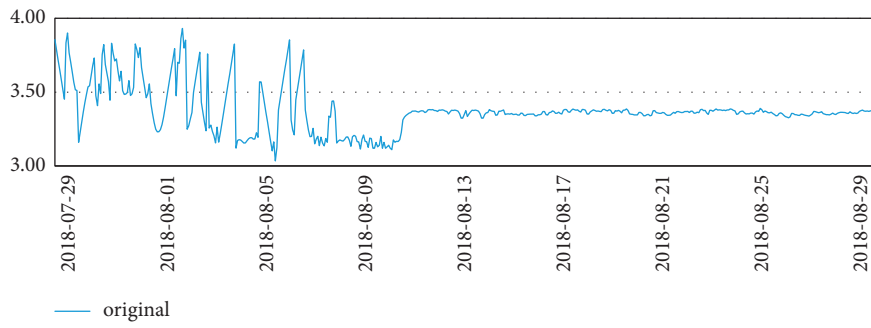


FIGURE 4: Temperature data.

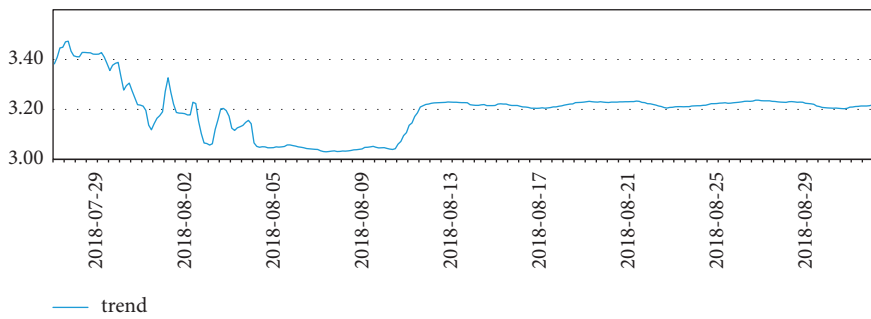


FIGURE 5: Temperature data.

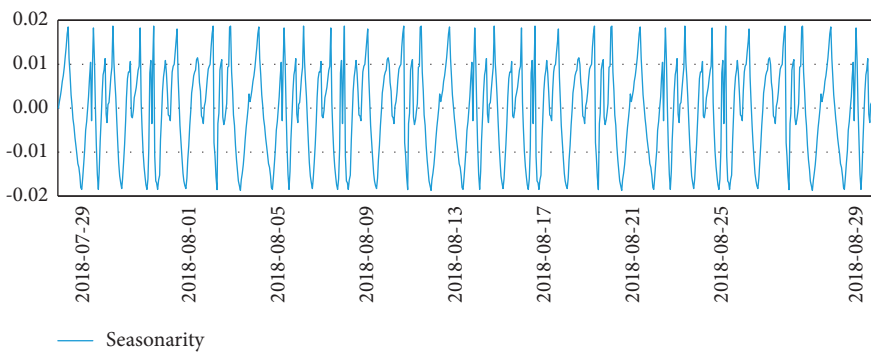


FIGURE 6: Temperature data.

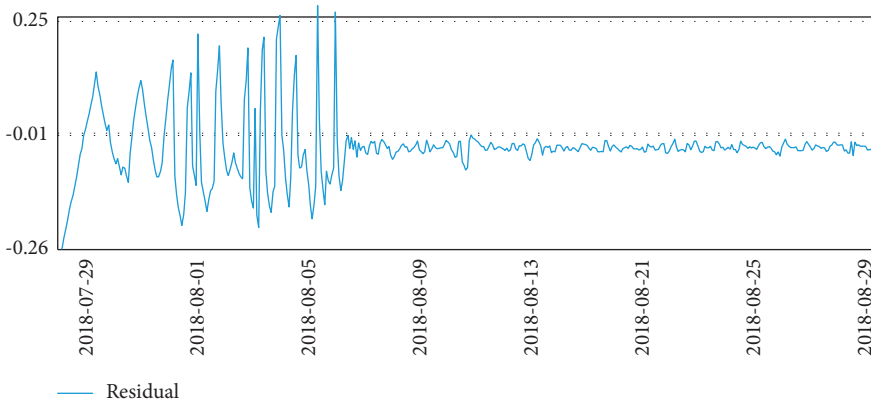


FIGURE 7: Temperature data.

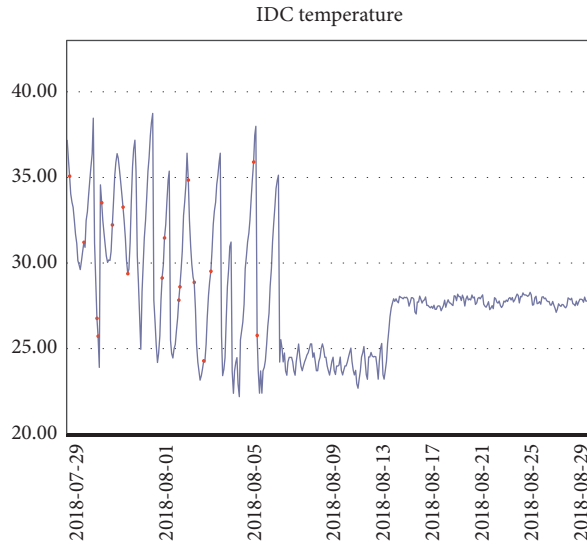


FIGURE 8: Dynamic threshold graph. ARIMA dynamics.

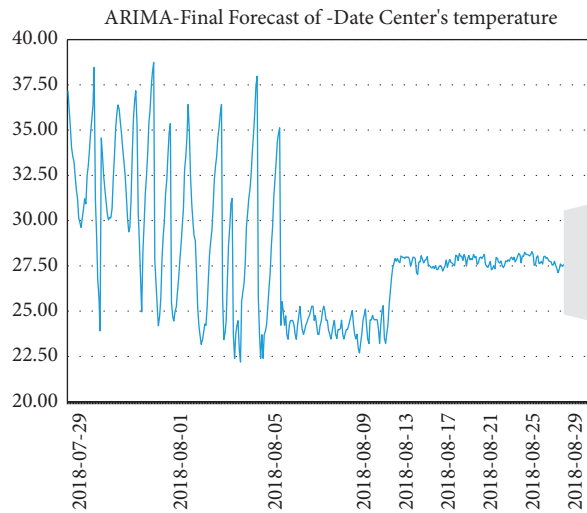


FIGURE 9: ARIMA prediction diagram.

interval is deducted from 2018/8/19 12:33 to calculate the time of the previous data. In this case, the step is 1 hour and the previous time is "2018/8/19 11:33." Then, the most recent data, "2018/8/19 11:29 23.1," will be used as the value at "2018/8/19 11:33," and so on.

Finally, the best combination of hyperparameters was obtained and the prediction was completed, but the processed data are needed to be reversed in order to recover the actual meaning of the data. The reason for this is the double difference operation performed on the initial data.

In terms of the technical specifications of the algorithms, RMSE and MAPE were mainly used to evaluate the performance of LSTM, ARIMA, and Holt-Winters. Interestingly, the top ranking is Holt-Winters rather than LSTM or ARIMA.

4.7. Results. ARIMA is not the best algorithm for predicting data center room temperature compared to Holt-Winters, but it performs better than LSTM.

Those red dots represent the anomalous data; the light blue background is the dynamic threshold based on the temperature values; the dark blue line is the real data; whenever the real data come in, this data are compared with its predicted value. Once the difference value exceeds the current threshold, an alarm is raised and the data are marked as a red exception in the graph and in the system, respectively, as shown in Figure 8.

As shown in Figure 9, the ARIMA prediction chart shows another feature, the prediction function, which can help to do some preparation before an event occurs. In the abovementioned figure, the green line on the far right with the grey background is the prediction line.

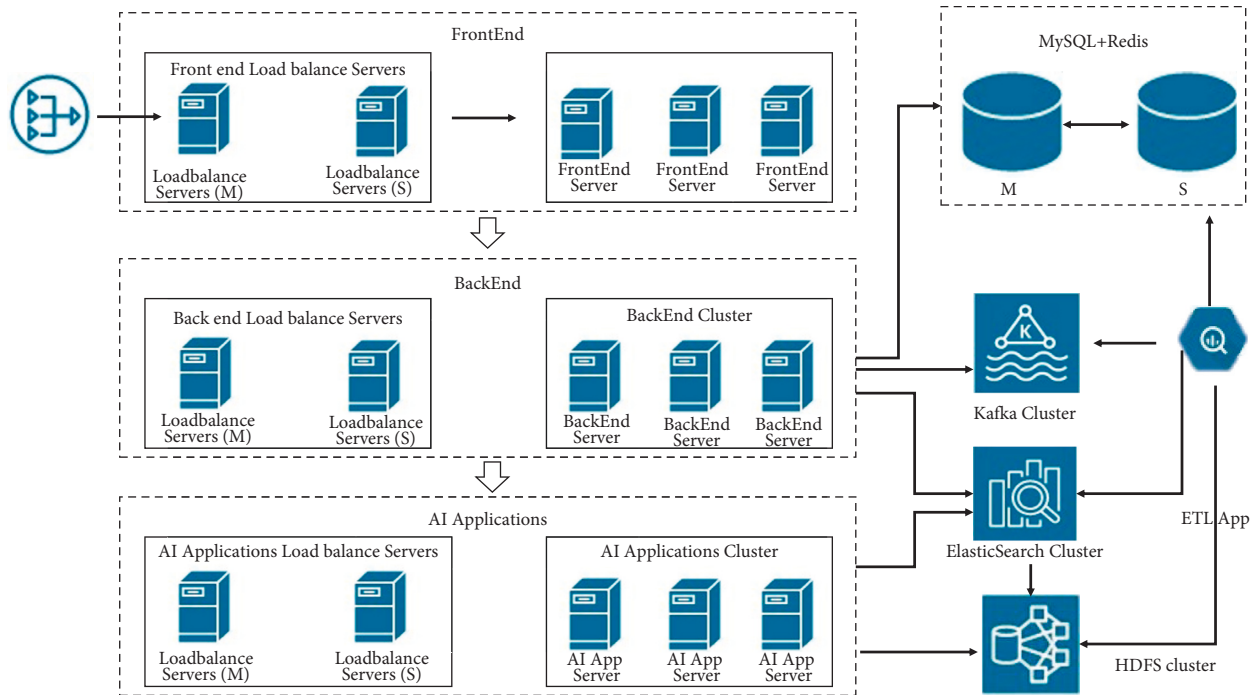


FIGURE 10: Standard deployment diagram.

Deployment architecture: typically, there are 3 types of deployment options, namely small, standard, and large.

Small deployment means that all the modules are deployed on a single server, which requires corresponding high-performance servers.

Here, the standard deployment diagram is given. Since the standard deployment is scalable, to get the large deployment, only the standard scheme needs to be extended, as shown in Figure 10.

5. Conclusion

Many real-world practices have shown that a working architecture or algorithm cannot be directly replicated from another domain due to completely different business scenarios and different data characteristics. In this paper, an AIOps architecture for the data center infrastructure monitoring domain is presented. The details of a proven working architecture and design in practice are presented in terms of core modules such as technical architecture, machine learning algorithms, big data, and business applications. Focusing on the technical part rather than every part of the architectural work, this paper presents a complete and detailed AIOps architecture. AIOps and APM (application performance management IT infrastructure management) systems are rapidly evolving; however, AIOps technology has not made any significant progress in data center infrastructure management. It is designed for data center on-site infrastructure monitoring. The mechanism and architecture design are described by the implementation and real-world performance of a number of different machine learning algorithms. Using data center room temperature data as an example, the test data clearly demonstrate the

effectiveness of this architecture. In addition, this architecture can be a useful reference for those who want to know how to implement machine learning algorithms in data center site monitoring systems.

- (1) This is a productive working architecture. As a time-series database, Elasticsearch is effective and can be used as a time-series database of choice.
- (2) The most suitable algorithm for the problem is probably the common one, not the most popular one.

In fact, these algorithms should be allowed to compete with each other by running them with the same dataset and the results will tell the winner. Therefore, the algorithm that is best suited to the data and scenario is the best algorithm.

In addition, the research in this paper has some limitations, mainly the following.

5.1. AI Governance Capability. As AI models continue to grow, there is a need to continuously monitor these models. Some metrics such as fairness, interpretability, robustness, and transparency should also be taken into account. On the other hand, AI systems themselves need to be governed. Specifically, engineering, development, deployment, and maintenance should be tracked and audited throughout the product lifecycle.

5.2. Multidimensional Relevance. In some cases, the accuracy and effectiveness of anomaly detection, especially those predicted by a single feature, are not always that good. Therefore, better results can be obtained if multiple correlated features could be used with other appropriate algorithms, e.g., correlation algorithms. In addition, it is often

necessary to identify those relevant features and complete the prediction in one algorithm, e.g., a neuronal network.

However, the problem of how to automatically select features can arise, especially when there are hundreds of features. A reliable approach is to use some automatic function or library to complete the first round of automatic selection. Then, these selected features can be verified by real data or expert experience. Finally, a reasonable feature set can be obtained.

5.3. *Service Network*. Different development languages and frameworks are used, such as Java, Vue, Scala, and Python, but no service grid or other microservice frameworks are used to unify these different services into one layer.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have conflicts of interest.

References

- [1] Y. Zhang, W. Yan, G. S. Hong et al., "Data fusion analysis in the powder-bed fusion AM process monitoring by Dempster-Shafer evidence theory," *Rapid Prototyping Journal*, vol. 28, no. 5, pp. 841–854, 2022.
- [2] T. Georghiou, E. Massou, S. Morris et al., "The impact of post-hospital remote monitoring of COVID-19 patients using pulse oximetry: a national observational study using hospital activity data," *EClinicalMedicine*, vol. 48, Article ID 101441, 2022.
- [3] P. J. Zhang, C. S. Wang, G. S. Wu, and Yu Wang, "Temperature gradient models of steel-concrete composite girder based on long-term monitoring data," *Journal of Constructional Steel Research*, vol. 194, 2022.
- [4] R. Shang, Z. Zhu, J. Zhang et al., "Near-real-time monitoring of land disturbance with harmonized Landsats 7–8 and Sentinel-2 data," *Remote Sensing of Environment*, vol. 278, 2022.
- [5] A. Heller, P. W. Nicholas, T. Brandon, W. B. Colby, and M. Z. Davidson, "An inductive approach to developing ecological site concepts with existing monitoring data," *Rangeland Ecology and Management*, vol. 83, 2022.
- [6] M. A. S. Rasha and R. R. Ali, "Monitoring of oil spill in the offshore zone of the Nile Delta using Sentinel data," *Marine Pollution Bulletin*, vol. 179, 2022.
- [7] F. S. Erica, J. R. Orin, R. B. Emily et al., "The potential of semi-structured citizen science data as a supplement for conservation decision-making: validating the performance of eBird against targeted avian monitoring efforts," *Biological Conservation*, vol. 270, 2022.
- [8] D. Rossat, J. Baroth, M. Briffaut et al., "Bayesian updating for nuclear containment buildings using both mechanical and hydraulic monitoring data," *Engineering Structures*, vol. 262, Article ID 114294, 2022.
- [9] J. Li, Z. Wu, J. Chen, X. Lu, and Z. Li, "FEM-Bayesian Kriging method for deformation field estimation of earth dams with limited monitoring data," *Computers and Geotechnics*, vol. 148, 2022.
- [10] S. Rusne, S. Arnout, B. Vasileios, K. Meister, W. Alexander, and V. T. Arjan, "European waste statistics data for a circular economy monitor: opportunities and limitations from the amsterdam metropolitan region," *Journal of Cleaner Production*, vol. 358, 2022.
- [11] T. Wu, L. Tang, S. Shao et al., "Accurate structural displacement monitoring by data fusion of a consumer-grade camera and accelerometers," *Engineering Structures*, vol. 262, 2022.
- [12] J. Hou, H. Jiang, C. Wan et al., "Deep learning and data augmentation based data imputation for structural health monitoring system in multi-sensor damaged state," *Measurement*, vol. 196, Article ID 111206, 2022.
- [13] A. Niranga and A. Giriraj, "Agricultural drought monitoring in Sri Lanka using multisource satellite data," *Advances in Space Research*, vol. 69, no. 11, 2022.
- [14] F. Huseynov, D. Hester, E. J. O'Brien et al., "Monitoring the condition of narrow bridges using data from rotation-based and strain-based bridge weigh-in-motion systems," *Journal of Bridge Engineering*, vol. 27, no. 7, 2022.
- [15] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: a survey," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2022.
- [16] S. M. Madzunkov, D. Nikolić, J. Simcic, A. Belousov, M. P. Gonzalez, and M. R. Darrach, "Data analysis and isotopic ratios measured onboard the Spacecraft Atmosphere Monitor," *International Journal of Mass Spectrometry*, vol. 477, 2022.
- [17] C. Jose, V. Laura, M. Cole et al., "Health monitoring of a conveyor belt system using machine vision and real-time sensor data," *CIRP Journal of Manufacturing Science and Technology*, vol. 38, pp. 38–50, 2022.
- [18] C. Pei, W. Zhao, S. Zhang, and Y. Wei, "Monitoring and early warning model based on multi-dimensional structure of power grid data," *Energy Reports*, vol. 8, no. S6, pp. 539–545, 2022.
- [19] P. E. Sang, Y. T. Jung, and H. C. Kim, "Monitoring permafrost changes in central Yakutia using optical and polarimetric SAR data," *Remote Sensing of Environment*, vol. 274, 2022.
- [20] I. Fuentes, P. José, and R. Vervoort, "Willem. Towards near real-time national-scale soil water content monitoring using data fusion as a downscaling alternative," *Journal of Hydrology*, vol. 609, 2022.
- [21] G. Zuzana and Z. Rehman, "Appraisal of the hypoplastic model for the numerical prediction of high-rise building settlement in Neogene clay based on real-scale monitoring data," *Journal of Building Engineering*, vol. 50, 2022.

Research Article

Deep-Learning-Based Financial Message Sentiment Classification in Business Management

Chen Shao ¹ and Xiaochen Chen²

¹*School of Intelligent Engineering, Shandong Management University, Jinan 250357, China*

²*School of Labor Relations, Shandong Management University, Jinan 250357, China*

Correspondence should be addressed to Chen Shao; shaochen@sdmu.edu.cn

Received 9 June 2022; Accepted 1 July 2022; Published 18 July 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Chen Shao and Xiaochen 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.

A deep-learning-based financial text sentiment classification method is proposed in this paper, which can provide a reference for business management. In the proposed method, domain adaptation is adopted to solve the common problem of insufficient labeled samples in the financial textual domain. Specifically, in the classification process, the seq2seq model is firstly adopted to extract the abstract from the financial message, which can reduce the influence of invalid information and speed up processing. In the process of sentiment classification, a bidirectional LSTM model is adopted for classification, which can more comprehensively make use of context information. Experiments are carried out to testify the proposed method through the open-source data set. It can be seen that the proposed method can effectively transfer from the reduced Amazon data set to the StockTwits financial text data set. Compared with the parameter-frozen-based method and the SDA-based method, the recognition rates have improved by 0.5% and 6.8%, respectively. If the target domain data set can be directly adopted for training, the recognition rate of the proposed method is higher than that of the SVM method and the LSTM method by 8.3% and 4.5%, respectively.

1. Introduction

Text-based sentiment classification has significant importance for business management. The effect of sentiment classification is mainly twofold. Firstly, through sentiment analysis of current financial texts, the current economic situation can be revealed, so as to adjust corresponding corporate strategies. Secondly, the comments on the corporate's products and the corporate itself can be performed for sentiment classification, so as to provide references for the corporate's products or the corporate itself.

Among the mentioned two aspects, sentiment analysis based on financial texts has a particularly significant effect on business management. For example, by extracting important features from financial texts, including financial news, financial comments, and financial-theme-related social networks, it is possible to perceive the emotions of the crowd and predict the financial or economic situation trends. This can also provide references for the strategic team, including

making the financing strategy, management strategy of the corporate, and so on. In this way, as information is obtained from a large amount of text data from a large crowd, the influence of subjective factors from the management personnel can be effectively eliminated. Therefore, sentiment analysis from financial texts has become a research hotspot in recent years.

Currently, there already exists extensive research for sentiment analysis of texts. There are mainly two types of methods. The first type is the traditional machine learning based methods, and the other is the deep-learning-based methods. However, no matter what type of method it is, effective mathematical representations of texts are needed at first. For early text representation methods, the texts are regarded as merely a set of words with the same weight, such as the bag-of-words (BoW) method [1–3], which has ignored the underlying grammar and sequences of words. Similar methods include the term frequency-inverse document frequency (TF-IDF) method [4–6], which goes a step further

and takes into account the different importance of different words. In this method, it is believed that the importance of words is proportional to the frequency of occurrence in the text and inversely proportional to the frequency of occurrence in the corpus. Currently, the most widely adopted method for text representation is the word2vec method [7–9]; the main idea is to use vectors to represent different words. For both the representational vector space and the real word meaning space, the measures of distance between words are similar. In the word2vec method, the research focus is on how to find a mapping from words to vectors so that the similarity of words in the vector space is preserved. The methods for establishing this mapping relationship include the continuous bag-of-words (CBOW) model [10, 11] and the skip-gram model [12, 13], which are both trained through the corpus to obtain the representational vectors of different words. In this context, this process is also referred to as a pretraining process, and currently, there already exists vector representations pretrained adopting a rich corpus.

For the first type of sentiment classification method mentioned above, the sentiment of texts is analyzed by adopting traditional classifiers, such as support vector machines (SVM), random forests, multinomial naïve Bayes (mNB), and other models. The related literature on these methods are [14–16]. In these traditional methods, the types of classifiers are different, while the features adopted are similar, which are not fully utilized, resulting in a limited recognition rate. For the second type of method, as deep-learning-based techniques can extract more abstract or high-level features from the text, often a better classification performance can be reached.

In this paper, by estimating the current economic situation through financial texts, business management can be improved. However, for such a purpose, the use of deep learning for sentiment classification from financial texts has two main problems: (1) the problem of insufficient amount of labeled financial texts and (2) the problem of establishing an efficient and effective model to improve the classification accuracy. Existing methods are then introduced following the logic of how to solve the mentioned two main problems.

In terms of sentiment analysis from financial texts, currently there exists an obvious problem, which is the lack of sufficient labeled financial texts. For this problem, a relatively straightforward solution is to adopt the labeled samples from other textual domains to make up for the lack of labeled financial textual samples. Then transfer learning can be applied for that. In literature [17], fine tuning is proposed for transfer learning from other textual domains to the financial textual domain. In this method, after a model is trained with a large number of labeled samples, the parameters of some layers are then frozen, mostly the shallow feature extraction layers in the front, and then other parameters are fine-tuned adopting a small number of labeled samples in the financial text domain. Noting that in this method, only the parameters of the abstract classification layer are fine-tuned to achieve transfer learning purposes. In this method, since the feature extraction parameters of the shallow layers are frozen, and new features cannot be

obtained for the classification of new data domains, the transfer learning performance is often limited. In [18], a combination of supervised and unsupervised methods is proposed. Specifically, the input samples are dimensionally reduced adopting the stacked denoising autoencoder (SDA) to obtain the shared feature vector among different input domains, and then the sentiment is estimated by another classifier. In this method, as the labeled samples are not adopted in the feature extraction process, the capability of this method is limited in terms of feature extraction. In general, although this method has comparatively good generality among different sample domains, its classification performance is limited. The authors in [19] have proposed an adversarial learning method to solve the problem of an insufficient number of labeled samples. In this adversarial network, there is a compromise between the loss of sentiment classification and the loss of domain classification. In [20], the idea of active learning is adopted, and the number of labeled samples required in the target domain in training are directly reduced by the method of active learning. However, in this active learning method, the reduction of labeled samples is only validated terms without active learning, and a large number of labeled samples are still needed.

In addition to solving the problem of insufficient labeled samples, it is also necessary to select a suitable model for textual sentiment classification. In [21], it is proposed that in order to better perform sentiment classification, multiple sources of information can be used, such as text information, images, or even videos. This is also referred to as multimodal classification. The method of multimodal classification can be considered as a type of methods for solving the problem of lack of training samples. The idea is to add more types of information other than textual information. However, to make it work, there still needs more training information. In literature [22, 23], the textual features and visual features are fused through the attention-based recurrent neural network (RNN) model and the adversarial network model, respectively, which can improve the classification accuracy compared to the single information source. In [24], it is proposed to further apply multimodal classification to the field of fake information detection. The two different feature vectors are merged through the multimodal compact bilinear pooling (MCBP) method, which can effectively fuse the two kinds of information for classification. In this paper, since most of the samples for financial sentiment analysis only contain textual information, only the unimodal sentiment classification (from textual information) is studied. Existing classification models can also be divided into two categories, convolutional neural network (CNN) based models and RNN-based models. The related literature include [25–29]. Generally speaking, among the two types, the RNN-based models generally have better classification performance due to better utilization of contextual information.

A deep-learning-based financial text sentiment classification method is proposed in this paper, which can provide a reference for business management. In the proposed method, domain adaptation is adopted to solve the common problem of insufficient labeled samples in the financial

textual domain. A subnetwork for distinguishing domains is added to the common sentiment classification network, and domain-related loss is also added to the overall cost function. In the proposed method, sentiment classification of financial text is possible without labeled financial text samples (i.e., target domain sample labels). In addition, in the classification process, the seq2seq model is firstly adopted to extract the abstract from the financial message, which can reduce the influence of invalid information and speed up the processing. In the process of sentiment classification, a bi-directional (long short-term memory) LSTM model is adopted for classification, which can more comprehensively make use of context information. In addition, an attention mechanism is also added to ensure that the information usage is comprehensive, while critical information are more weighted. Experiments are carried out to testify the proposed method through the open-source data set. It can be seen that the proposed method can effectively transfer from the reduced Amazon data set to the StockTwits financial text data set. Compared with the parameter freezing method and the SDA-based method, the recognition rates have improved by 0.5% and 6.8%, respectively. If the target domain data set can be directly adopted for training, the recognition rate of the proposed method is higher than that of the SVM method and the LSTM method by 8.3% and 4.5%, respectively.

2. Methods

The flow chart of the proposed method in this paper is shown in Figure 1. It can be seen that the proposed method can be divided into three parts. The first part is to adopt the seq2seq model to extract the abstract of the message. Here, the extracted abstract is of fixed length. Moreover, the seq2seq model adopted in this paper is not trained in the implementation of the method but is already pretrained by adopting existing rich corpus. In our method, it is directly put to use as a general abstract extraction model. Therefore, the impacts of different textual domains on abstract extraction have not been taken into consideration here. After the abstract extraction, a highly summative abstract of the text is obtained, which can reduce the influence of invalid information on classification. The second part of the proposed method contains a feature extraction subnetwork based on the bidirectional LSTM model, which can use the extracted abstract information (also the output of the seq2seq model) to further summarize the abstract. The output is a feature vector, which can be used for both sentiment classification and the domain classification. There are two domains here: the source domain and the target domain. The source domain refers to a large number of texts (mainly product comments in our implementation) with sentiment classification labels, and the target domain refers to financial texts, which is not labeled. The third part contains two paralleled classification subnetworks, the sentiment classification subnetwork and the domain classification subnetwork. The sentiment classification subnetwork is adopted for the final text sentiment classification, and the sentiment classification cost is

constructed from the labeled textual samples. The domain classification subnetwork is adopted to distinguish sample domains, with the domain discrimination cost function. The proposed method in this paper is introduced as follows according to the three parts mentioned above. Noting that here the domain adaptation can be considered as a framework of the proposed framework. With different losses, including the sentiment classification loss and the domain classification loss added to the proposed method, domain adaptation is realized. However, for the seq2seq-model-based abstraction part, it is a standalone processing flowing. In our implementation, the extracted abstract is of fixed length. Moreover, the seq2seq model adopted in this paper is not trained in the implementation of the method but is already pretrained by adopting existing rich corpus.

2.1. Abstract Extraction Based on the Seq2seq Model. The seq2seq model adopted in this paper is described in Figure 2. The seq2seq model is often adopted in the field of automatic translation in natural language processing (NLP), where the input is of one language and the output is another language. In our paper, the pretrained seq2seq abstract extraction model is directly adopted; the input is the descriptive text; and the output is a fixed-length abstract. Since the application of the seq2seq model in abstract extraction has already been widely studied, only some key points are described herein.

As shown in Figure 2, the basic structure of the seq2seq model is encoder-decoder-based. The context vector is the connection between the encoder and the decoder, which is the output of the encoder and the input of the decoder. This vector can be regarded as an effective summary of the input according to the encoder, which contains the contextual information of the input sequence. Generally speaking, both the encoder and decoder contain long short-term memory (LSTM), so as to keep the memory of the sequence. However, the pure encoder-decoder structure can suffer from the loss of sequence information. When the input sequence is long and the dimension of the context vector is low, more information can be neglected from the encoder. This can lead the decoder to work abnormal due to the loss of key information.

To solve the problem, a straightforward idea is to add different importance to the input sequence. Therefore, attention-based encoder-decoder model is proposed in [30]. In the attention-based model, the context variable is the weighted average of the encoder output hidden variables h_j and weights w_{ij} , which can be expressed as follows:

$$cv_i = \sum_{j=1}^T w_{ij} h_j. \quad (1)$$

The weights of different hidden variables w_{ij} are calculated according to e_{ij} . e_{ij} is jointly determined by the hidden variables of the encoder and the decoder at different times (where h_j denote the output hidden variables of the encoder and c_j denote the output hidden variables of the decoder) according to the score function $\text{get_score}(\cdot)$. The score

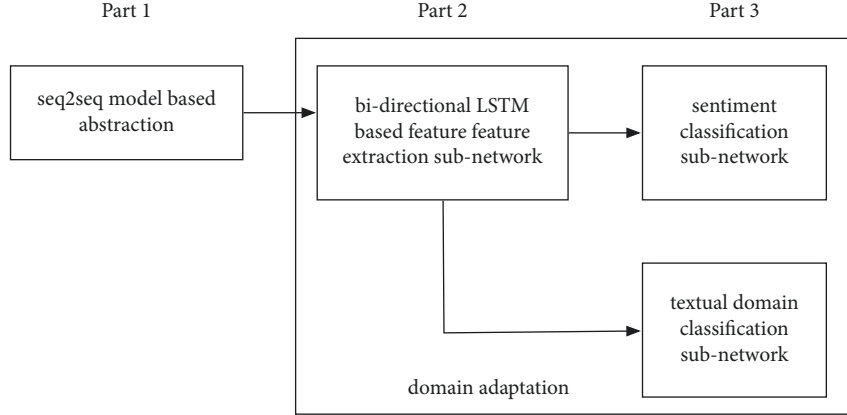


FIGURE 1: The processing flowchart of the proposed method in three parts.

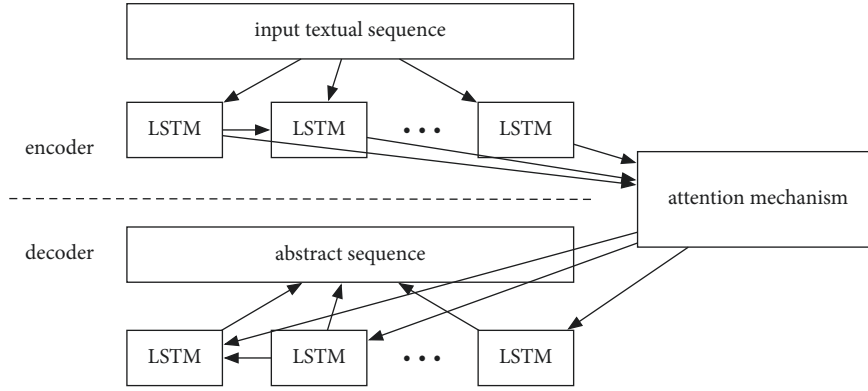


FIGURE 2: The encoder-decoder structure of the adopted seq2seq model.

function is also called the alignment function, and the dot product of the two hidden vectors is directly applied in this paper.

$$w_{ij} = \frac{\exp e_{ij}}{\sum_{k=1}^T \exp e_{ik}},$$

$$e_{ij} = \text{get_score}(c_j, h_j), \quad (2)$$

$$\text{get_score}(\cdot) \equiv (\cdot).$$

In the seq2seq model, the output is a series of sequences and their respective probabilities. Usually, in order to avoid the probability of accidental output, only one sequence is finally taken as the maximum probability output, and the BeamSearch algorithm can be used to obtain the sequence output with the maximum overall probability.

2.2. Feature Extraction Subnetwork Based on Bidirectional LSTM. The proposed feature extraction subnetwork based on bidirectional LSTM is applied to extract the information in the abstract and to generate the corresponding feature vector, which can be adopted for the subsequent sentiment classification and the domain classification. Note that here the feature extraction subnetwork is different from a common feature extraction network. It not only needs to

take into account the extraction of sentiment classification-related features but also the features that can distinguish source or target domain data samples. Its basic structure is shown in Figure 3.

According to the descriptions in the paper [31], the model of bidirectional LSTM can extract more comprehensive features because the extracted feature is related to the total semantic information. By adding an attention mechanism on this basis, it can be ensured that the key information is retained in a focused manner as well as the comprehensiveness of the information. In Figure 3, F represents the final state of the bidirectional LSTM model, that is, the addition result of the state values of the forward and backward hidden layers.

The calculation process of the attention probability distribution and the calculation process of the final feature vector with the attention mechanism are as follows, which generally include two main steps:

$$a_n = \frac{\exp(h'_n)}{\sum_{i=1}^N \exp(h'_i)}, \quad (3)$$

$$h'_n = h_n^T U F,$$

where a_n denotes the attention probability distribution of the hidden layer unit at time n , h'_i denotes the input

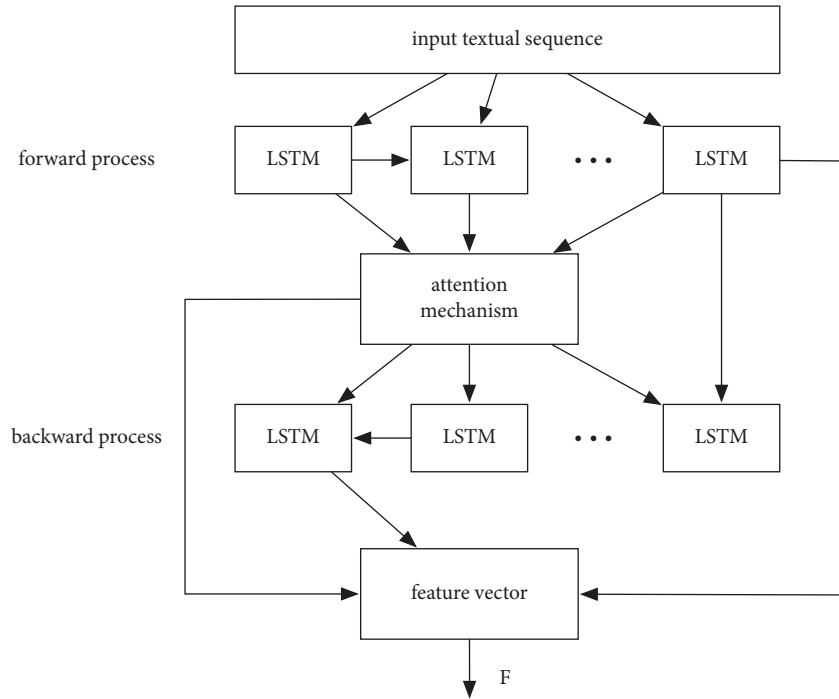


FIGURE 3: The feature extraction subnetwork based on the bidirectional LSTM model.

sequence, and U denotes the weighting matrix. h_n denotes the sum of the bidirectional hidden layer state values at the time n . With the bidirectional LSTM model, the feature vector F can be obtained. The difference between the abstract extraction model and the feature extraction model is as follows. For the abstract extraction process, the model is just based on a one direction seq2seq model with an encoder-decoder structure. However, for the feature extraction process, the model is bidirectional LSTM-based. The model of bidirectional LSTM can extract more comprehensive features because the extracted feature is related to the total semantic information.

2.3. Sentiment Classification Subnetwork and Domain Classification Subnetwork. After the feature vector is obtained, it can be adopted for sentiment classification and domain classification by the sentiment classification subnetwork and the domain classification subnetwork, respectively. In this paper, the sentiments in the financial text domain are bullish and bearish, respectively, and the different domains are the financial text domain and other data domains, respectively.

In this paper, the structures of the sentiment classification subnetwork and the domain classification subnetwork are shown in Figures 4 and 5, respectively. It can be seen that the two subnetworks have similar structures. The only difference is that the domain classification subnetwork has an additional layer of reverse gradient layer in the beginning, which will be introduced later. The mutual structure in both the sentiment classification subnetwork and the domain subnetwork is introduced as follows, including a fully connected layer and a softmax layer. The operation of the softmax layer is as follows:

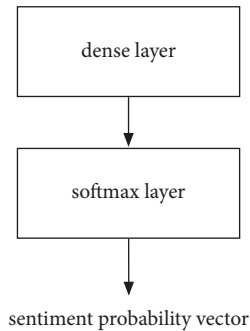


FIGURE 4: The structure of the sentiment classification subnetwork.

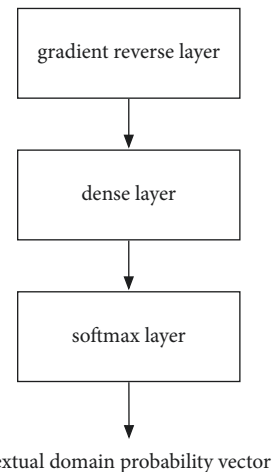


FIGURE 5: The structure of the domain classification subnetwork.

$$l_i = \frac{\sigma^i}{\sum_{j=1}^N \sigma^j}, \quad (4)$$

where σ^i represents the component of the i th dimension in the input vector, l_i denotes the output probability of the i th dimensional component, and N represents the total number of classes.

The purpose of the domain classification subnetwork is to distinguish whether the data samples originated from the financial textual domain. The addition of the reverse gradient layer mentioned above can achieve the following effects: (1) enabling the domain classification subnetwork to distinguish whether it originated from the financial textual domain, (2) while making the features in the feature extraction subnetwork unable to extract features that are sensitive to specific domains. The mentioned effects can achieve the purpose of transfer learning. The inverse gradient layer can be expressed as follows:

$$\begin{aligned} \mathbf{F}(\mathbf{s}) &= \mathbf{s}, \\ \frac{\partial \mathbf{F}(\mathbf{s})}{\partial \mathbf{s}} &= -\eta \mathbf{I}, \end{aligned} \quad (5)$$

where the function $\mathbf{F}(\cdot)$ denotes the function of forward propagation and \mathbf{s} denotes the input vector. In the forward propagation process of the reverse gradient layer, the layer has no special operation on the input data sample. However, during training, the gradient of back propagation is reversed and linearly reduced in this layer. The addition of the reverse gradient layer has enabled domain classification subnetwork training while making the feature extraction subnetwork unable to extract features related to domain classification. Therefore, the purpose of transfer learning can be achieved.

The proposed network has two cost functions in the training process, namely the sentiment classification cost function and the domain classification cost function. The two cost functions have similar structures, and they are both cost functions based on cross-entropy, which are

$$\begin{aligned} \text{Loss}_{1,i}(\cdot) &= -\sum d_{1,i} \log l_{1,i}, \\ \text{Loss}_{2,j}(\cdot) &= -\sum d_{2,j} \log l_{2,j}, \end{aligned} \quad (6)$$

where the subscript 1 is sentiment classification related and 2 is domain classification related; $\text{Loss}(\cdot)$ denotes the cost function; l denotes the corresponding output through the softmax layer, that is, the probability of the corresponding label; and d denotes the real label (i.e., the real label vector representation after one-hot encoding). The subscripts i and j represent the index of the training samples.

Then the final cost function can be expressed as the sum of two subcost functions as follows:

$$\text{Loss}_{\text{all}}(\cdot) = \text{Loss}_1(\cdot) - \eta \text{Loss}_2(\cdot) \quad (7)$$

A minus sign and a linear parameter η are added to the domain classification cost function, which is equivalent to the operation of the reverse gradient layer mentioned above. Setting the cost function in this way is equivalent to only adding the reverse operation of the gradient during the

training process, without affecting the forward inference process.

3. Results and Discussion

As transfer learning is needed in this paper, so data sets of source and target domains are both required. To validate the proposed method in this paper, the selected data sets the source and target domains are both open-sourced. The selected source domain is the classic reduced Amazon data set. The data set comprises comments of four products: books, DVDs, electronics, and kitchen appliances. Each category contains 2,000 labeled reviews, including 1,000 positive reviews and 1,000 negative reviews, each accounting for 50%. We chose the financial social network StockTwits as the target domain data set. The labels of this data set are bullish and bearish. In order to facilitate domain adaptation, we correspond bullish to positive labels in the reduced Amazon data set and bearish to negative labels. Although we have not used the labels of the target domain during training, the labels of the target domain are actually known and can be adopted for statistical purposes. For the above data sets, the training set and the test set are split according to 70% and 30%.

Since the proposed method can be divided into two aspects, the first is based on transfer learning between different data domains; the second is abstract extraction by seq2seq; and then the recognition is based on bidirectional LSTM. Therefore, here the experimental comparison is also set up according to these two aspects.

3.1. Comparison of Transfer Learning Methods. For comparisons of transfer learning methods, the SDA-based method and the parameter-frozen-based method described above are selected for comparison. Note that among these methods, the proposed method in this paper and the SDA-based method can both be regarded as an unsupervised transfer learning method, which does not require any labeled samples from the target domain during training. The method of parameter frozen is semisupervised, which requires a small number of labeled samples from the target domain.

The comparison results of the three methods are shown in Figure 6 and Table 1. It can be seen that the accuracy of domain classification of the proposed method has reached 61.2%. Compared with the parameter-frozen-based method and the SDA-based method, the recognition rate has improved by 0.5% and 6.8%, respectively. The results have fully illustrated the effectiveness of the proposed method in textual domain classification, which can provide an effective solution for the problem of the small number of labeled samples in the target domain mentioned above. Although the improvement is not obvious compared with the parameter frozen method, it should be noted that the parameter frozen method has adopted a small number of labeled samples in the target domain, while the proposed method has not. Noting that the results here are already the overall results. The comparisons in the next part are made

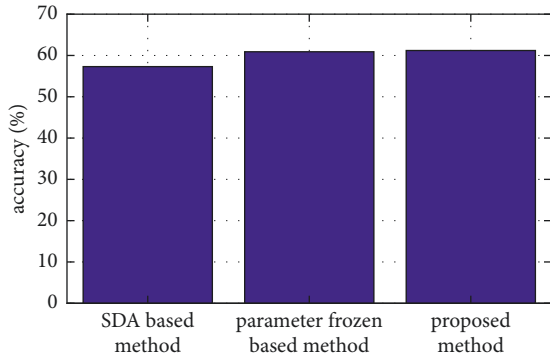


FIGURE 6: The histogram of the average recognition rate of different methods.

TABLE 1: Comparison of average recognition rates of different methods.

Method	Accuracy (%)	Improvement (%)
Proposed method	61.2	—
SDA-based method	57.3	6.8
Parameter-frozen-based method	60.9	0.5

without considering the domain adaptation framework so that the effects of other factors including the abstract extraction and the bidirectional LSTM model can be tested.

3.2. Comparison of Classification Methods. After the transfer learning part is verified, the recognition effect of the proposed seq2seq abstract extraction + bidirectional LSTM feature extraction framework will be tested. In order to be able to compare with other recognition methods, in this section, we have skipped the source domain and have directly taken the labeled samples in the target domain for training. Here, the proposed method in this paper is also compared with the two recognition methods mentioned above. The first method is a traditional SVM-based method without deep learning. The second method is based on LSTM, which does not include the abstract extraction part and directly adopts the original message as the input. The comparison results are shown in Figure 7 and Table 2. It can be seen that the recognition rate of the proposed method in this paper has reached 80.7%. Compared with the SVM method and LSTM method, it has improved by 8.3% and 4.5%, respectively.

In order to further study the characteristics of the proposed method, the contribution of the abstract extraction and bidirectional LSTM factors to the final recognition accuracy is also looked into. There are three cases for comparison: (1) adding abstract extraction and bidirectional LSTM feature extraction (the proposed), (2) without abstract extraction, only with bidirectional LSTM feature extraction, (3) with abstract extraction and with ordinary LSTM-based feature extraction, and (4) without abstract extraction and with ordinary LSTM-based feature extraction. Table 3 has

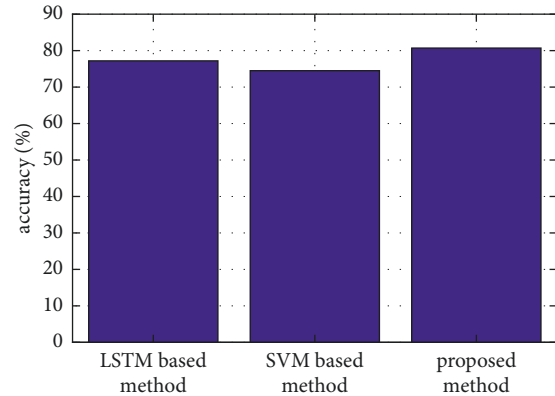


FIGURE 7: Comparison of the classification accuracy of the proposed method, the traditional SVM method, and the LSTM-based method.

TABLE 2: Comparison of the classification accuracy of the proposed method, the traditional SVM method, and the LSTM-based method.

Method	Accuracy (%)	Improvement (%)
Proposed method	80.7	—
SVM-based method	74.5	8.3
LSTM-based method	77.2	4.5

TABLE 3: The comparisons of the proposed method and the other three cases.

Situation	Accuracy (%)	Improvement (%)
Case 1	80.7	—
Case 2	77.9	3.6
Case 3	78.2	3.2
Case 4	77.2	4.5

listed the different recognition rates in the four cases. It can be seen that the factors of abstract extraction and bidirectional LSTM both have a significant impact on improving recognition accuracy. Compared with cases 2, 3, and 4, the recognition accuracy of case 1 has improved by 3.6%, 3.2%, and 4.5%, respectively.

4. Conclusions

A deep-learning-based financial text sentiment classification method is proposed in this paper, which can provide a reference for business management. The proposed method has adopted the domain adaptation framework to solve the common problem of insufficient labeled samples in financial text classification. Specifically, in the classification process, the seq2seq model is first adopted to extract the abstract from the financial message. Then, a bidirectional LSTM model is adopted for classification, which can more comprehensively make use of context information. Experiments are carried out to testify the proposed method through the open-source data set. It can be seen that the proposed

method can effectively transfer from the reduced Amazon data set to the StockTwits financial text data set. Compared with the parameter-frozen-based method and the SDA-based method, the recognition rates have improved by 0.5% and 6.8%, respectively. If the target domain data set can be directly adopted for training, the recognition rate of the proposed method is higher than that of the SVM method and the LSTM method by 8.3% and 4.5%, respectively. In addition, results have also shown that the factors of adding abstract extraction and adopting the bidirectional LSTM model can both be effective in improving performance.

Data Availability

The data adopted in the paper are available at <https://stocktwits.com>.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] Y. Zhang, R. Jin, and Z.-H. Zhou, "Understanding bag-of-words model: a statistical framework," *International Journal of Machine Learning and Cybernetics*, vol. 1, no. 1-4, pp. 43–52, 2010.
- [2] H. M. Wallach, "Topic modeling: beyond bag-of-words," in *Proceedings of the 23rd international conference on Machine learning*, New York, NY, USA, June 2006.
- [3] L. Lei Wu, S. C. H. Hoi, and N. Yu, "Semantics-preserving bag-of-words models and applications," *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1908–1920, 2010.
- [4] J. Ramos, "Using tf-idf to determine word relevance in document queries," *Proceedings of the first instructional conference on machine learning*, vol. 242, no. 1, 2003.
- [5] Ho C. Wu, R. W. P. Luk, K. F. Wong, and K. L. Kwok, "Interpreting tf-idf term weights as making relevance decisions," *ACM Transactions on Information Systems*, vol. 26, no. 3, pp. 1–37, 2008.
- [6] B. Trstenjak, S. Mikac, and D. Donko, "KNN with TF-IDF based framework for text categorization," *Procedia Engineering*, vol. 69, pp. 1356–1364, 2014.
- [7] Y. Goldberg and O. Levy, "word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method," 2014, <https://arxiv.org/abs/1402.3722>.
- [8] W. Ling, C. Dyer, A. W. Black, and I. Trancoso, "Two/too simple adaptations of word2vec for syntax problems," in *Proceedings of the 2015 conference of the North American chapter of the association for computational linguistics: human language technologies*, Denver, CO, USA, May 2015.
- [9] K. W. Church, "Word2Vec," *Natural Language Engineering*, vol. 23, no. 1, pp. 155–162, 2017.
- [10] R. Xie, Z. Liu, J. Jia, H. Luan, and M. Sun, "Representation learning of knowledge graphs with entity descriptions," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, 2016.
- [11] Qi Wang, J. Xu, H. Chen, and B. He, "Two improved continuous bag-of-word models," in *Proceedings of the 2017 international joint conference on neural networks (IJCNN)* IEEE, Anchorage, AK, USA, May 2017.
- [12] D. Guthrie, and B. Allison, W. Liu, L. Guthrie, and Y. Wilks, "A closer look at skip-gram modelling," *LREC*, vol. 6, 2006.
- [13] Y. Song, S. Shi, J. Li, and H. Zhang, "Directional skip-gram: explicitly distinguishing left and right context for word embeddings," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 2, New Orleans, LA, USA, June 2018.
- [14] E. Hossain, S. Sharif, and M. Moshikul Hoque, "Sentiment polarity detection on Bengali book reviews using multinomial naïve bayes," *Advances in Intelligent Systems and Computing. Progress in Advanced Computing and Intelligent Engineering*, pp. 281–292, Springer, Singapore, 2021.
- [15] N. Chirawichitchai, "Sentiment classification by a hybrid method of greedy search and multinomial naïve bayes algorithm," in *Proceedings of the 2013 Eleventh international conference on ICT and knowledge engineering*, IEEE, Bangkok, Thailand, November 2013.
- [16] P. Karthika, R. Murugeswari, and R. Manoranjithem, "Sentiment analysis of social media network using random forest algorithm," in *Proceedings of the 2019 IEEE international conference on intelligent techniques in control, optimization and signal processing (INCOS)*, April 2019.
- [17] J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," 2018, <https://arxiv.org/abs/1801.06146>.
- [18] M. Yang, W. Yin, Q. Qu, W. Tu, Y. Shen, and X. Chen, "Neural attentive network for cross-domain aspect-level sentiment classification," *IEEE Transactions on Affective Computing*, vol. 12, no. 3, pp. 761–775, 2021.
- [19] Y. Zhang, J. Li, H. Wang, and S. C. T. Choi, "Sentiment-guided adversarial learning for stock price prediction," *Frontiers in Applied Mathematics and Statistics*, vol. 7, p. 8, 2021.
- [20] S. Li, Y. Xue, Z. Wang, and G. Zhou, "Active learning for cross-domain sentiment classification," in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, New York, NY, USA, August 2013.
- [21] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. I. Satoh, "Spotfake: a multi-modal framework for fake news detection," in *Proceedings of the 2019 IEEE fifth international conference on multimedia big data (BigMM)*, pp. 39–47, IEEE, Singapore, September 2019.
- [22] S. Singhal, A. Kabra, M. Sharma, R. R. Shah, T. Chakraborty, and P. Kumaraguru, "SpotFake+: a multimodal framework for fake news detection via transfer learning (student abstract)," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 10, Article ID 13915, May 2020, New York, NY, USA.
- [23] L. Cui, S. Wang, and D. Lee, "Same: sentiment-aware multi-modal embedding for detecting fake news," in *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining*, pp. 41–48, Vancouver, Canada, August 2019.
- [24] A. Fukui, D. H. Park, D. Yang, A. Rohrbach, T. Darrell, and M. Rohrbach, "Multimodal compact bilinear pooling for visual question answering and visual grounding," 2016, <https://arxiv.org/abs/1606.01847>.
- [25] F. Yu, Q. Liu, S. Wu, L. Wang, and T. Tan, "A convolutional approach for misinformation identification," in *Proceedings of IJCAI, Melbourne Australia*, August 2017.
- [26] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions" in *Proceedings of the IEEE conference on computer vision and pattern recognition*, Boston, MA, October 2015.
- [27] J. Ma, W. Gao, Z. Wei, Y. Lu, and K. F. Wong, "Detect rumors using time series of social context information on

- microblogging websites,” in *Proceedings of the 24th ACM international on conference on information and knowledge management*, New York, NY, USA, October 2015.
- [28] P. Bahad, P. Saxena, and R. Kamal, “Fake news detection using bi-directional LSTM-recurrent neural network,” *Procedia Computer Science*, vol. 165, pp. 74–82, 2019.
- [29] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi, and B. W. On, “Fake news stance detection using deep learning architecture (CNN-LSTM),” *IEEE Access*, vol. 8, Article ID 156695, 2020.
- [30] H. Bahuleyan, L. Mou, O. Vechtomova, and P. Poupart, “Variational attention for sequence-to-sequence models,” 2017, <https://arxiv.org/abs/1712.08207>.
- [31] H. Yang, J. Zhang, S. Li, and T. Luo, “Bi-direction hierarchical LSTM with spatial-temporal attention for action recognition,” *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 1, pp. 775–786, 2019.

Research Article

Application of Data Mining Technology in Enterprise Green Innovation Model Construction and Path Analysis

Binfeng He 

Jinling College, Nanjing University, Nanjing 210089, Jiangsu, China

Correspondence should be addressed to Binfeng He; 031109001@jlxj.nju.edu.cn

Received 12 June 2022; Accepted 29 June 2022; Published 13 July 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Binfeng He. 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.

Since sustainable development has become the dominant mode of human development at the present stage, green technology has received more and more attention under this background. The development of green technology has become an important means to achieve sustainable development. Green technological innovation is a kind of technological innovation. However, because the goal of green technological innovation is different from that of traditional technological innovation, the dynamic mechanism of green technological innovation lies in the similarities and differences of traditional technological innovation. This paper focuses on data mining technology to design and optimize the enterprise green technology creation model. At present, clustering algorithm and association rule algorithm are important research contents in big data mining technology. Among them, the clustering algorithm refers to the process of grouping similar data objects in a large amount of data information, so that the approximate data information can be aggregated and clustered, which is convenient for data mining calculation. In the algorithm, the shortcomings of the original clustering algorithm, such as insufficient data processing and incomplete analysis, are improved, and the data processing is improved by 51.7%, which has a good processing effect on subsequent data preprocessing and dynamic incremental clustering. In the follow-up experiment, the role of green technology in corporate finance is reflected.

1. Introduction

Greening is a concept. When this concept is introduced into the field of production and technological innovation and gradually transformed into practice, it will improve people's welfare and change our lives. The greening of technology and technological innovation is to achieve this goal which is a powerful tool [1]. The principle of circular economy is reduction, reuse, and recycling, which is characterized by low exploitation of resources, high utilization, and low emission of pollutants. Obviously, these are all unable or difficult to be strongly supported by the existing conventional technology [2]. We must take green technology as the guarantee, accelerate the development and demonstration of resource conservation and alternative technologies with universal popularization significance, energy cascade utilization technology, extension of industrial chain and related industrial link technology, zero emission technology, alternative technology of toxic and harmful raw materials,

recycling and treatment technology of renewable resources, green remanufacturing and other technologies, and strive to break through the technical obstacles restricting the development of circular economy [3]. Under this strong pursuit of a better environment, technological innovation, especially the development of green technology innovation, is bound to be promoted [4]. And how does green technology innovation affect the environment, and how we should take various specific measures to conform to the trend of social development and establish a good innovation power mechanism in order to enhance the green technology innovation capability of the whole society, which is worth studying the problem.

Like human beings, an enterprise is a natural entity and a part of nature at first. No matter how far it develops, it cannot be separated from nature, and it must obtain resources from nature [5]. Therefore, the sustainable development of enterprises must be based on the continuous access to resources [6]. Green technology innovation is an

effective means to implement the strategy of sustainable development. Green technology innovation meeting the requirements of environmental development can improve resource utilization, save energy and raw materials, reduce environmental pollution, reduce environmental externality loss in production and consumption, and improve the ability of enterprises to internalize environmental costs. Of course, this is mainly reflected in the long term [7]. From a global perspective, the extensive model at the cost of energy consumption, environmental damage and ecological service function decline not only creates a miracle of economic growth, but also brings huge ecological environment pressure to economic development and social progress. The aggravation of environmental pollution and the imbalance of resource income distribution lead to the crisis of sustainable development [8]. At present, the country is vigorously promoting a new model of energy conservation and pollution reduction, and developing green industries and energy conservation and environmental protection. Green economic growth has become a strategic choice for my country's economy in the future. For this reason, in the research of this paper, based on the perspective of green development, how to support green development through technological innovation and practice the concept of ecological environmental protection is expounded [9]. Profit maximization is the eternal theme of enterprise development. Especially in today's era of knowledge economy, with the rapid development of economic globalization and network information technology, scientific and technological progress is changing with each passing day, and market competition is becoming increasingly fierce. Enterprises must have sustained competitiveness if they want to develop continuously.

Innovation is the source for enterprises to obtain competitiveness. As a main form of information, data is undoubtedly the main carrier of Internet Communication [10]. Advances in information technology have also been accompanied by an explosion of data. The application of computer has brought unparalleled changes to people's way of life, work, and study. In traditional society, all the behaviors people do in real life can gradually be realized on the Internet. Data mining technology is a technology that analyzes information that is closely related to people's lives from seemingly unconnected data. Data mining technology involves many fields, including database theory in computer technology and artificial intelligence theory in automation technology [11]. Generally speaking, the analysis methods of data mining mainly include classification data mining, clustering data mining, association rule data mining, and outlier data mining. Aiming at various business problems, different analysis methods will get more accurate conclusions. The uses of data mining mainly include but are not limited to these five aspects: classification data mining, clustering data mining, association rule data mining, prediction data mining, and deviation data mining. These five functions do not exist independently. They will affect each other in the project of data mining. Data mining technology generally includes decision tree method and artificial neural network method. However, the above-mentioned research

has not well constructed the innovative model design and improvement of enterprise green technology based on data mining. Therefore, this paper proposes the following innovations:

- ① Based on the high efficiency and information processing ability of data mining technology, according to the nature of green technology, this paper will combine data mining and analysis technology to build an enterprise green technology system model. As modern enterprises have huge data in technology, how to analyze and process data has become the key work of enterprises. Data mining technology is superior in dealing with complicated data, and it integrates management association rules in basic technology.
- ② Because the clustering method is often used in data processing, but because the clustering and association algorithm in the current data mining technology has low efficiency for data processing and unreasonable algorithm structure, this paper designs and improves the data mining algorithm to achieve high efficiency and easy processing in attribute weighting and density, so as to get convenience and efficiency in practical processing.

The chapters of this paper are arranged as follows. Section 1 is the introduction part, which discusses the background and significance of the topic selection of the paper and expounds the innovation points of the paper. Section 2 of this paper mainly combines the domestic and foreign research results of data mining technology in the field of green technology and proposes innovative results and research ideas of this paper. Section 3 of this paper is the method part, which deeply discusses the application and principle of related algorithms and puts forward a new green technology data analysis and processing model based on the previous research results and the innovation of this paper. Section 4 of this paper mainly discusses the experimental part of the algorithm application. Through the experimental results, on the basis of sorting out the data, the enterprise green technology model is established. Section 5 is the conclusion, which summarizes the research results and related work.

2. Related Work

Chanmee and Kesorn believed that green technology innovation is also called ecological technology innovation, which belongs to a kind of technological innovation. Generally, management innovation and technological innovation aiming at protecting the environment are collectively referred to as green technology innovation. From the perspective of product life cycle, green technology innovation refers to the whole process of green technology from the formation of ideas to the market. It is an innovation to reduce the cost of product life cycle [12]. Lin et al. research shows that the retrieval results of users applying traditional techniques cannot be complete relative to their retrieval goals, while the retrieval results of different users with the

same retrieval goals have greater coverage and integrity than the retrieval results of each user, so the cooperation between users will produce overall benefits [13]. Zhang believes that because the dynamic mechanism of technological innovation is to study the driving factors of technological innovation and how the driving factors affect the technological innovation behavior of enterprises, without effectively solving the problem of the dynamic mechanism of technological innovation, the research on other issues will lack solid theories basis [14]. Saeed believes that statistics and analysis of market and customer data through data mining technology, acquisition of customer knowledge, discovery of market opportunities, determination of target customer groups and marketing mix, and scientific formulation of market and product strategies have become more and more concerned by modern enterprises [15]. The research of Martinelli F, mercaldo F, Nardone V, and others shows that as a main form of information, data is undoubtedly the main carrier of Internet communication. The progress of information technology is also accompanied by explosive data. Over the years since the birth of the Internet, the amount of data in various fields of human society, especially in the field of information, has accelerated, resulting in the concept of big data [16]. Li et al. believe that in the field of artificial intelligence, data mining has been replaced by knowledge discovery, another professional term. In the database field, the term data mining is more recognized, so data mining is also known as database knowledge discovery [17]. The research results of Chavets L O, Chahovets V V, and Didenko A S show that from a technical point of view, data mining can be regarded as a collection of a series of advanced algorithms, mathematical models, and systems [18]. Through the use of various data mining algorithms, the interpretation of potential classification relations, clustering relations, and regular relations of data can be realized, so as to realize high mining of data value, and help decision makers in various industries to make more scientific and reasonable decisions according to the interpretation and prediction of these data. Qi et al. believe that due to the massive, high-dimensional, heterogeneous, dynamic, spatio-temporal, diversity, multi-source, multi-scale, fuzziness, and other characteristics of big data, the internal correlation between data is very hidden, so it is necessary to deeply explore its potential value and internal correlation through association rule algorithm [19]. According to Du and Zhao, the use of data mining is mainly but not limited to these five aspects: classification data mining, clustering data mining, association rule data mining, prediction data mining, and deviation (singularity) data mining. These five functions do not exist independently, and will influence each other in the data mining engineering [20]. Gefen and Larsen think that CRM lacks the management and application of “knowledge owned by customers.” They introduce knowledge management technology into CRM framework, and combine the processes of knowledge acquisition, storage, sharing and use, and put forward a framework of CRM analysis system based on knowledge management [21]. Zhang and Li believe that the successful integration of industrialization and informatization can realize the transformation of production

mode, promote the change of economic development mode, promote the adjustment and upgrading of industrial structure, improve the ability of technological innovation, and finally realize the rapid and healthy development of the whole national economy [22]. Sang H uses the theoretical principles and policies of technological innovation and ecological civilization construction to conduct research on how to promote green development through technological innovation and progress and practice the concept of environmental protection. This has played a certain role in promoting the concept of sustainable development in my country and practicing the value theory of ecological civilization has a positive effect [23]. Geng et al.’s green economy emphasizes the word “green.” His main representative works are new energy, and the following have developed low-carbon related industries, low-carbon technologies, green products beneficial to mankind, new finance characterized by green investment, and green grid system [24].

Based on the research of the above related work, this paper determines the positive role of data mining technology in the field of enterprise green technology, constructs a financial management model based on the combination of various technologies, makes in-depth analysis and research on the obtained and collected data using data mining technology, makes more effective use of data, and mines the valuable information hidden behind the data, so as to simplify and improve the enterprise green technology.

3. Methodology

3.1. Relevant Theoretical Research and Analysis

3.1.1. Data Mining Technology. The essence of data mining is to perform algorithm operations from a large amount of noisy, uncertain, and fuzzy real business data, and finally discover data knowledge that has not yet been recognized or cannot be clearly recognized and has certain practical meanings in the process. Among them, the representative works are: discovering rules and distinguishing rules in relational database by attribute-oriented induction method; association rules are found in the transaction database; clustering analysis and optimization based on distance and density. In addition, decision tree, neural network, genetic algorithm, rough set, fuzzy set, and visualization methods have also been studied and applied. At present, there are many kinds of clustering algorithms and classification algorithms, including different data mining algorithms. In practical application, the selection and use of specific algorithms are mainly determined according to the object objectives, so as to achieve the predetermined data analysis and mining results. Broadly speaking, a complete data mining process should consist of six steps: data collection, data selection, data preprocessing, establishment of data mining models, and data mining operations on target data using the established data mining models to explain and express the data mining results. In the whole process of data mining, data preparation is the foundation work, data mining implementation is the core step, and ending expression and explanation is the necessary evaluation process.

Because the results produced in the whole process cannot meet the initial set problems, it is necessary to reprocess the whole process. Therefore, it is the normal state in data mining to iteratively carry out these three steps. Figure 1 shows the basic flow chart of data mining.

The connotation understanding of the concept of technological innovation has a process of continuous development and deepening. Scientific and technological innovation is the process of applying scientific discovery and technological invention to the production process and improving product value and competitiveness. Technological innovation is the product of the coevolution of the triple helix structure of scientific discovery, technological invention, and market application. Scientific and technological innovation is all scientific and technological activities and economic activities that create new knowledge, generate new technologies, and apply new knowledge and new technologies throughout the entire scientific and technological activities. The classification algorithm is a data mining algorithm that classifies the data set to discover the correlation and difference of a large number of sample data, so as to deeply mine the value of the data. Taking the research of decision tree algorithm at home and abroad as an example, the earlier decision tree algorithm is ID3 algorithm, and its timeliness of data classification is poor. Technological innovation capability is a series of capability elements involved in the process of adapting to changing market demand, researching and developing new products to meet market demand through accumulated knowledge and experience, and providing products to consumers after commercial manufacturing and successful marketing.

3.1.2. Enterprise Green Technology. Green technology is also known as environmentally friendly technology or ecological technology. The content it covers involves a general term for technologies, processes or products that reduce environmental pollution and reduce the use of raw materials, natural resources, and energy. This concept originated from the reflection on the situation that modern technology destroys the ecological environment and threatens human existence. It can be regarded as one of the signs of ecological philosophy, ecological culture, and even ecological civilization. The key to sustainable green innovation of enterprises is motivation. Only by solving the power problem of enterprise's green sustainable innovation can enterprises actively cultivate their own green sustainable innovation ability, and can they seriously solve a series of other problems in the operation of green sustainable innovation. Green technology innovation is an effective means to implement the strategy of sustainable development. Green technology innovation meeting the requirements of environmental development can improve resource utilization, save energy and raw materials, reduce environmental pollution, reduce environmental externality loss in production and consumption, and improve the ability of enterprises to internalize environmental costs. Of course, this is mainly reflected in the long term. The influencing factors of green technology innovation are mostly studied from the internal and external

aspects. It is generally believed that the internal factors mainly include: the pursuit of profit, the improvement of competitiveness, the enhancement of corporate value, and the convenience of financing; the external factors mainly include: social demand, government policy orientation, technology promotion, and market competition incentives. Figure 2 shows the basic model of technological innovation process.

From the above discussion on the technological innovation process model of the comprehensive role of technology and market, it can be found that compared with the early single factor driven model, the upgrading of the model is inseparable from the theoretical research and practice of technological innovation. The above-mentioned research and innovation process model of technology and enterprise green technology is a more representative and typical innovation model. This model completely reflects the process of technological innovation and divides it into different but relatively independent influences. Although the functions are not necessarily continuous in stages, they are logically connected, reflecting that technological innovation not only requires a general technical process, but also requires corresponding organization and management, reflecting that the technological innovation process is a fusion of technical capabilities and market needs.

3.2. Analysis of the Elements of Enterprise Green Technology Innovation. For the special technological innovation of green technological innovation, a single element cannot explain the driving force of green technological innovation. The driving force of green sustainable innovation of enterprises mainly comes from the continuous pursuit of maximizing economic and social benefits and the goal of long-term development of enterprises. Green technology has a special nature. It strives to obey the rules and norms of nature and adapt itself to various habits of human beings. The technology system based on green technology will greatly alleviate the distortion and destruction of the relationship between man and nature caused by traditional modern technology system, and effectively overcome the anti-ecological characteristics of modern technology. The external factors of the enterprise include the promotion of science and technology, the change of market demand, and the change and development of market mechanism, socio-economic culture, ecological environment, and political system. All these factors are coupled, developed, and interacted with each other, forming a strong and lasting systematic driving force for the sustainable green innovation of the enterprise. Figure 3 shows the competitiveness of enterprise green technology.

From the above figure, we know that green consumption has become the mainstream. Besides product quality, price and after-sales service, greenness is another key indicator for consumers to choose products or services. In the era of green economy sweeping across the country and abroad, it is an important task for modern enterprises how to carry out clean production with no pollution or low pollution, how to meet consumers' growing green demand by reducing costs

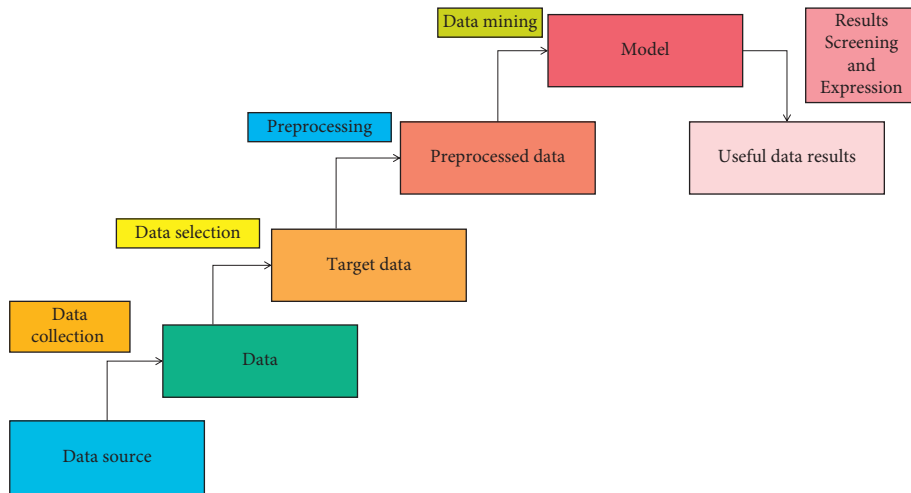


FIGURE 1: Basic flow chart of data mining.

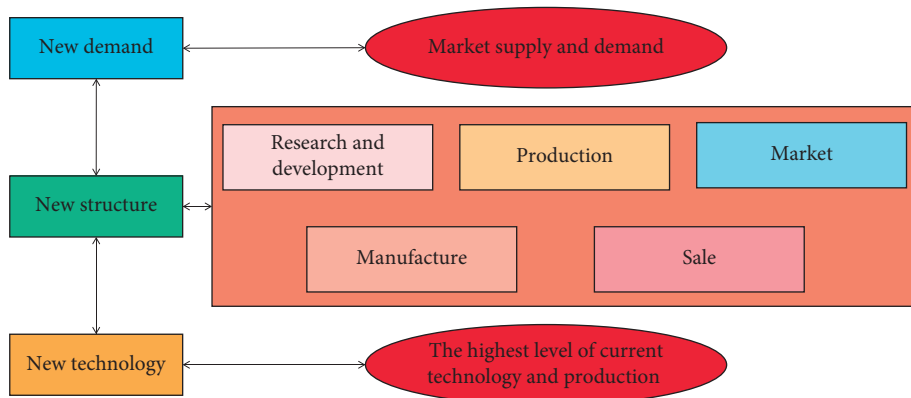


FIGURE 2: Basic model diagram of the technological innovation process.

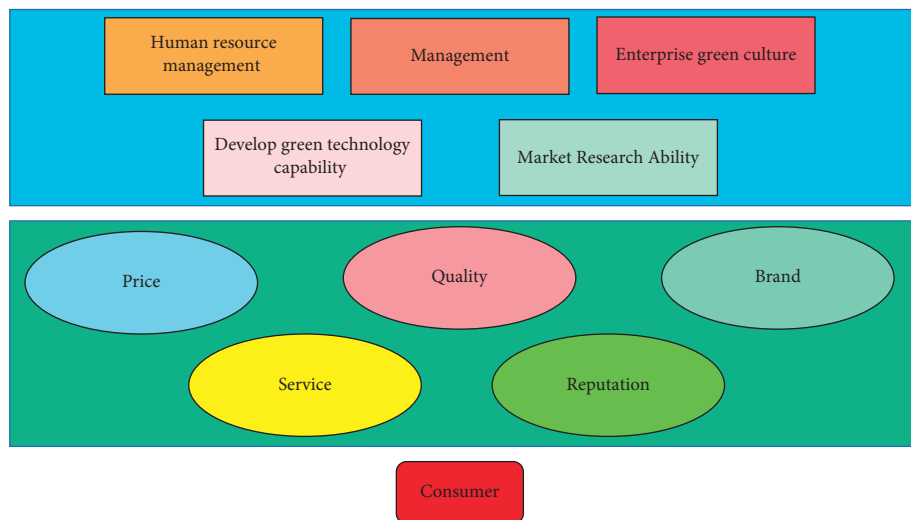


FIGURE 3: Schematic diagram of the competitiveness of enterprises' green technology.

and consumption, and how to obtain higher profits than other competitors. At the same time, systems are not closed to each other, but open. Materials, information, and energy are exchanged between the systems. The internal cause plays

a decisive role in the development of things, and the external cause always works through the internal cause. The market external economy of green technology innovation reflects the contradiction between green technology innovation

pursuing the maximization of ecological benefits and industrial activities pursuing the maximization of economic benefits, the contradiction between the nature of public goods and the nature of private goods in green technology and products, and the contradiction between the individualization of input and the popularization of the beneficiaries of the created eco-environmental value.

3.3. Basic Design and Optimization of Model Algorithm.

The improved association algorithm proposed in this paper can be mainly divided into the following two aspects of association rule definition: on the one hand, one or a pair of quantitative attributes is clustered to form qualified clusters or intervals; on the other hand, it searches for frequent clusters and obtains quantitative association rules based on the distance. It generally includes clustering features and association cluster features. Therefore, the following attribute information is defined for the above features:

Clustering characteristics: a summary of information describing subclusters of objects, including cluster information projected on other attribute sets. The formula is

$$CF(C_X) = \left(N, \sum_{i=1}^N t_i[X], \sum_{i=1}^N t_i[X]^2 \right). \quad (1)$$

In the above formula, N represents the number of all tuples in the subcluster. Associate cluster features. Assuming $C_X = t_1, t_2, \dots, t_n$, the formula is as follows:

$$ACF(C_X) = \left(N, \sum_{i=1}^N t_i[Y], \sum_{i=1}^N t_i[Y]^2 \right). \quad (2)$$

For the study of clusters, this paper also proposes an evaluation and analysis of density. Let $S[X]$ be the set of N data sets t_1, t_2, \dots, t_N projected to attribute set X , then the distance metric formula of $S[X]$ is

$$d(S[X]) = \frac{\sum_{i=1}^N \sum_{j=1}^N \delta_X(t_i[X], t_j[X])}{N(N-1)}. \quad (3)$$

In the above formula, δ represents the distance measure between tuples. The greater the distance of $S[X]$, the greater the deviation of its data set projection to the attribute set X . Therefore, it is also necessary to introduce interval division to isolate the density interval to facilitate data processing. A cluster X on attribute set C should be less than or equal to the density threshold d_0^X , and greater than or equal to the frequency value s_0 , and its formula is expressed as

$$\begin{aligned} d(C_X[X]) &\leq d_0^X, \\ C_X &\geq s_0. \end{aligned} \quad (4)$$

In the above formula, s_0 defines the minimum number of tuples in a cluster. When the data clusters satisfying the density distribution are denser, the data clusters satisfying the frequency threshold will have sufficient support. Because clustering considers the relative distance between data points or intervals, the range of quantitative attributes can be divided into appropriate partitions according to the value of

quantitative attributes, which effectively solves the problem of data set segmentation.

The information gain is an important indicator used to evaluate the ability of an attribute to discriminate the training data. The information gain of a C_i attribute can be obtained by the following mathematical methods. Let D be the training set of class label tuples, class label attributes have n different values, n different classes C_i ($i = 1, 2, \dots, n$), $C_i D$ is the set of tuples of classes in D , $|D|$ and $|C_i D|$ are the number of tuples in D and $C_i D$, respectively. Information gain is actually used to measure attribute selection in ID3 algorithm. It selects the attribute with the highest information gain as the splitting attribute of the node. This attribute minimizes the amount of information required for tuple classification in result partition. The desired information needed to classify the tuples in D can be obtained by the formula:

$$\text{Info}(D) = - \sum_{i=1}^n P_i \log P_i, \quad (5)$$

$$XO_i = \frac{\sum_{j=1}^{N_i} t_j^i[X]}{N_i}.$$

Now, suppose that the tuple in D is divided according to attribute A , and attribute A divides D into v different categories. Therefore, after the division, the following calculations are needed to get accurate measurement information in the future.

$$\text{Info}_A(D) = - \sum_{j=1}^v \frac{|D_j|}{|D|} * \text{Info}(D). \quad (6)$$

Therefore, according to the definition of information gain:

$$\text{Gain}(A) = \text{Info}(D) - \text{Info}_A(D). \quad (7)$$

By comparing the information gain of each attribute, the attribute with the largest information gain can be found. ID3 algorithm has obvious advantages. The algorithm is based on information theory, has simple logic process, and strong learning ability; But at the same time, the algorithm also has its corresponding disadvantages. The disadvantage of the algorithm is that when the data set is small, the algorithm is more effective, and the tolerance of the algorithm to noise is relatively low. Therefore, when the training data set becomes larger or the proportion of noise data increases, the final data mining results may be affected.

For the problem of production input in the actual green technology research and development of enterprises, this paper also reflects it in the algorithm part. Because enterprises invest in new technologies and pay attention to income and output, the production function is calculated by the following formula:

$$Y = F(C, T). \quad (8)$$

Among them, Y is the output quantity, C is the capital invested in production, and T is the technical level. In the

actual calculation, the algorithm can save 60% of the cost of the enterprise. When the technical level is within a reasonable range, it has a good effect. Since the clustering algorithm is the basic condition of the association analysis of the original algorithm, its clustering effect will directly affect the rationality of the interval division of each data set, and also have a significant impact on the effectiveness of the association analysis of the algorithm. Therefore, when clustering data, not only one or a pair of attributes, but also the clustering algorithm should be used to cluster according to all attributes of data objects, and the interval division or clustering should be formed according to the clustering results. Through the above optimization design, the green technology innovation model of enterprises has been greatly improved, and the efficiency of data processing has increased by 75.8%.

4. Result Analysis and Discussion

Based on the above introduction and the construction of the model design, this paper conducted a simulation experiment to verify and test the practicability and accuracy of the model, and judged the optimization and analysis advantages of the model on the enterprise green technology model through the experimental data and the processing of the experimental data. Figures 4 and 5 show the comparison of algorithms on the data set before and after optimization.

According to the comparison of the line graphs, in the optimization design of this paper, the simulation and density of the dynamic data set can be represented in the clustering process. Therefore, under a certain data density, due to the correlation between various data sets, the density attraction points are found in turn by calculating the density formula. Through comparison, it is found that the clustering algorithm is improved by 51.7% in data set processing, and it has a good processing effect in subsequent data preprocessing and dynamic incremental clustering. Figures 6 and 7 show the comparison of variance and standard deviation of data before and after optimization.

The above figure is the standard deviation and variance diagram of the data, which illustrates the standard deviation and variance of the data within the range of estimated values. It is almost the same as that before optimization, but it is more effective than that before optimization in terms of explaining the performance of the model. As mentioned earlier, the distribution of data is very close to the normal distribution, so it is relatively easy to reveal the “reliability” of model estimates according to the properties of standard deviation. For example, since 78% of the cases fall within plus or minus one standard deviation of the mean, in the worst case (when the estimated value is 0.5 in the figure), it can be deduced that 78% of the actual value will fall within plus or minus 0.32 centered on the estimated value (0.32 is derived from the maximum value of the figure matching the difference). Therefore, it can be said that when the predicted value of the model is 0.5, the actual value is in the range of 0.38 to 0.82 in 85% of cases. Similarly, for any point on the line, the reliability of its estimate can be described by how much of the actual value (expressed as a percentage) is

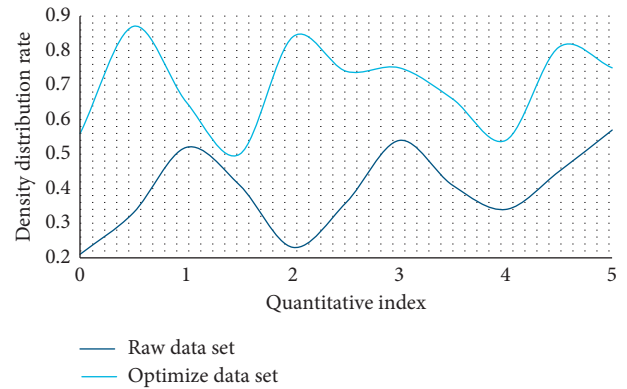


FIGURE 4: Density comparison between the original data set and the optimized data set.

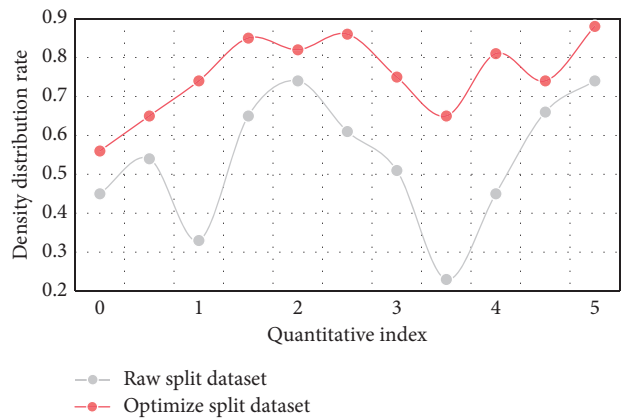


FIGURE 5: Comparison of segmentation data set before and after optimization.

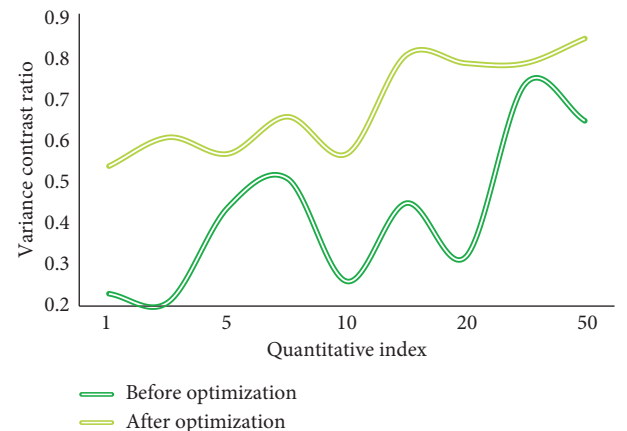


FIGURE 6: Comparison of data variance before and after optimization.

expected to fall within a certain range of distance estimates. Figure 8 shows the financial situation of the company before and after entering the green technology stage.

Through the above comparison, in the research and development and application of green technology, enterprises invested a lot of money in the research and development of new technologies in the early stage, but as the

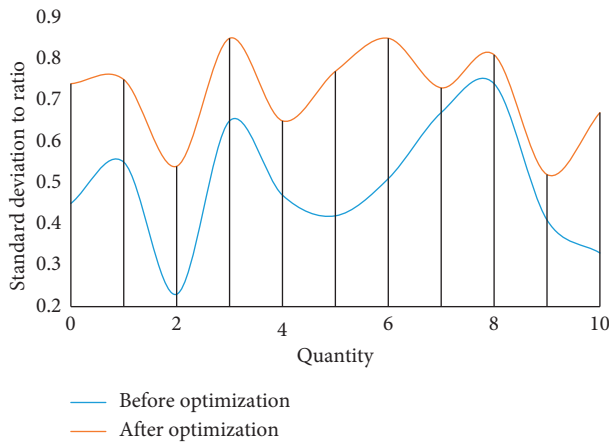


FIGURE 7: Comparison of standard deviations before and after optimization.

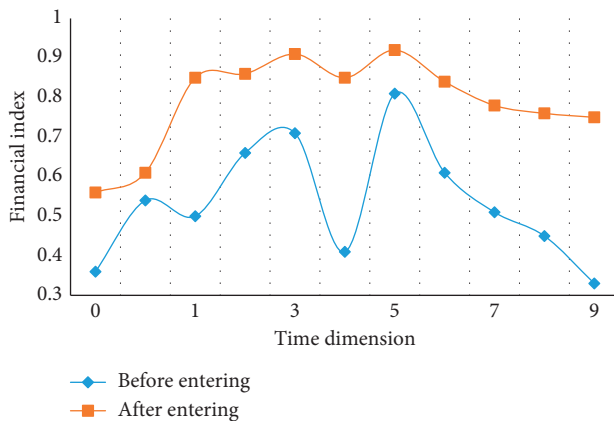


FIGURE 8: Comparison of the support of green technology to enterprises.

time line goes on, the proportion of green technology in the main competitiveness of enterprises gradually increases. It provides a strong technical support and basic guarantee for the sustainable development of enterprises. Although it requires great investment from enterprises in the early stage, it is absolutely necessary for enterprises with long-term plans to increase green technological innovation. For enterprise management, it is also an inevitable choice for enterprises to develop well in the future, and it can promote the optimization of enterprise financial structure and technological innovation.

5. Conclusions

With the rapid development of economic globalization and network information technology, scientific, and technological progress is changing with each passing day, and market competition is becoming increasingly fierce. Enterprises must have sustainable competitiveness if they want to develop continuously. Innovation is the source for enterprises to obtain competitiveness. Enterprise green sustainable innovation is an extremely complex process with the interaction of multiple internal and external factors. The root

of technological innovation power is changing with the times. Although the root of technological innovation in the market sense has always been the pursuit of profit, it is not completely the case today when the scientific concept of development is advocated. The way of sustainable development requires enterprises to implement green production and carry out green innovation. However, Chinese enterprises are full of confusion about how to implement green technological innovation, and the original technological innovation power model has been unable to meet the needs of the times. In order to further improve the efficiency of existing data mining algorithms and improve the results of data mining, this paper designs an efficient data mining algorithm for enterprise green technology based on the defects and problems of existing data mining algorithms. In addition, in the design, the density attraction points are found successively by calculating the density formula. Through comparison, it is found that due to the optimization and improvement in the clustering algorithm, the data set processing is improved by 51.7%, which has a good processing effect in the preprocessing of subsequent data and dynamic incremental clustering.

Data Availability

The data set can be obtained from the author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This work received the philosophy and social science research project of Jiangsu University: “Research on the key driving mechanism of innovation activities of small and medium-sized enterprises in Jiangsu Province-Analysis Based on the survey questionnaire of small and medium-sized enterprises in Jiangsu Province” (project approval no. 2021SJA2223) and the open topic of Jiangsu Productivity Society: “Research on the path and Countermeasures for the high-quality development of small and medium-sized enterprises in Jiangsu Province-Survey and analysis of small and medium-sized enterprises based on the perspective of technological innovation drive” (Project no. JSSCL2021B023).

References

- [1] R. Zheng and G. Zheng, “An artificial intelligence data mining technology based evaluation model of education on political and ideological strategy of students,” *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 5, pp. 1–12, 2020.
- [2] G. Ozkaya, M. Timor, and C. Erdin, “Science, technology and innovation policy indicators and comparisons of countries through a hybrid model of data mining and MCDM methods,” *Sustainability*, vol. 13, no. 2, pp. 1–52, 2021.
- [3] Y. Wu, X. Mu, and D. Kahrizi, “Design of biological gene information collection system based on data mining technology,” *Cellular and Molecular Biology (Noisy-Le-Grand, France)*, vol. 66, no. 7, pp. 93–102, 2020.

- [4] A. Malik, B. Pandey, and C. C. Wu, "Secure model to generate path map for vehicles in unusual road incidents using association rule based mining in VANET," *Journal of Electronic Science and Technology*, vol. 16, no. 2, pp. 59–68, 2018.
- [5] T. Deniz, I. B. Tulu, and T. Klemettai, "Analysis of ARMPS2010 database with La Model and an updated abutment angle equation," *International Journal of Mining Science and Technology*, vol. 30, no. 1, pp. 107–114, 2020.
- [6] Y. Zeng, "Evaluation of physical education teaching quality in colleges based on the hybrid technology of data mining and hidden markov model," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, p. 4, 2020.
- [7] L. Huang, "Design of an IoT DDoS attack prediction system based on data mining technology," *The Journal of Supercomputing*, vol. 78, no. 8, pp. 1–23, 2021.
- [8] C. Yuan, Y. Yang, and Y. Liu, "Sports decision-making model based on data mining and neural network," *Neural Computing and Applications*, vol. 2020, no. 2, pp. 1–14, 2020.
- [9] Z. J. Kuang, H. Zhou, D. D. Zhou, J. p. Zhou, and K. Yang, "A non-group parallel frequent pattern mining algorithm based on conditional patterns," *Frontiers of Information Technology & Electronic Engineering*, vol. 20, no. 9, pp. 1234–1245, 2019.
- [10] J. Zhou, "Research on the competence model of innovative entrepreneurial team by using data mining technology," *Revista de la Facultad de Ingenieria*, vol. 32, no. 4, pp. 741–748, 2017.
- [11] T. Liu, A. Shen, X. Hu, G. Tong, and W Gu, "The application of collaborative business intelligence technology in the hospital SPD logistics management model," *Iranian Journal of Public Health*, vol. 46, no. 6, pp. 744–754, 2017.
- [12] S. Chanmee and K. Kesorn, "Data quality enhancement for decision tree algorithm using knowledge-based model," *Current Journal of Applied Science and Technology*, vol. 20, no. 2, pp. 259–277, 2020.
- [13] Z. Lin, L. Yan, R. Wang, X. Fu, and Q. Lin, "Efficient privacy-preserving classification construction model with differential privacy technology," *Systems Engineering and Electronic Technology: English*, vol. 28, no. 1, p. 9, 2017.
- [14] J. Zhang, "Research on adaptive recommendation algorithm for big data mining based on Hadoop platform," *International Journal of Internet Protocol Technology*, vol. 12, no. 4, p. 213, 2019.
- [15] T. Saeed, "Data mining for small and medium enterprises: a conceptual model for adaptation," *Intelligent Information Management (English)*, vol. 12, pp. 183–197, 2020.
- [16] F. Martinelli, F. Mercaldo, V. Nardone, A. Orlando, A. Santone, and G Vaglini, "Model checking based approach for compliance checking," *Information Technology and Control*, vol. 48, no. 2, pp. 278–298, 2019.
- [17] Y. Li, J. Ge, and W. Hao, "Construction and simulation of a multiattribute training data mining model for basketball players based on big data," *Wireless Communications and Mobile Computing*, vol. 2021, no. 7, Article ID 6399266, 14 pages, 2021.
- [18] L. O. Chagovets, V. V. Chahovets, and A. S. Didenko, "The data mining technology applications for modeling the unevenness of socio-economic development of regions," *Business Inform*, vol. 3, no. 506, pp. 82–91, 2020.
- [19] Z. X. Qi, H. Z. Wang, and A. J. Wang, "Impacts of dirty data on classification and clustering models: an experimental evaluation," *Journal of Computer Science and Technology*, vol. 36, no. 4, pp. 806–821, 2021.
- [20] Y. Du and T. Zhao, "Network teaching technology based on big data mining and information fusion," *Security and Communication Networks*, vol. 2021, no. 9, Article ID 6629563, 9 pages, 2021.
- [21] D. Gefen and K. Larsen, "Controlling for lexical closeness in survey research: a demonstration on the technology acceptance model," *Journal of the Association for Information Systems*, vol. 18, no. 10, pp. 727–757, 2017.
- [22] J. Zhang and Y. Li, "Research on the application of data mining technology in the employment and entrepreneurship guidance of university students," *C e Ca*, vol. 42, no. 3, pp. 1059–1063, 2017.
- [23] H. . Sang, "Analysis and research of psychological education based on data mining technology," *Security and Communication Networks*, vol. 2021, no. 7, Article ID 8979507, 8 pages, 2021.
- [24] C. Geng, Y. Xu, and N. Metawa, "Intelligent financial decision support system based on data mining," *Journal of Intelligent & Fuzzy Systems*, vol. 2021, no. 2, pp. 1–10, 2021.

Research Article

Deep Learning-Based Football Player Detection in Videos

Tianyi Wang  and Tongyan Li

College of Physical Education, Qiqihar University, Qiqihar 161000, China

Correspondence should be addressed to Tianyi Wang; 02305@qqhru.edu.cn

Received 8 June 2022; Accepted 28 June 2022; Published 12 July 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Tianyi Wang and Tongyan 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.

The main task of football video analysis is to detect and track players. In this work, we propose a deep convolutional neural network-based football video analysis algorithm. This algorithm aims to detect the football player in real time. First, five convolution blocks were used to extract a feature map of football players with different spatial resolution. Then, features from different levels are combined together with weighted parameters to improve detection accuracy and adapt the model to input images with various resolutions and qualities. Moreover, this algorithm can be extended to a framework for detecting players in any other sports. The experimental results assure the effectiveness of our algorithm.

1. Introduction

In recent years, computer science has demonstrated great potential in the sport fields. For example, computer vision-based virtual reality was used for sports posture correction, [1] a computer vision-driven evaluation system was adopted for decision-making in sports training, [2] and object detection was used in sports analysis [3]. Sports analysis is crucial for improving athletes' performance. A classical method puts sensors to athletes' key positions and record raw data. Then data science methods are used to analyze the data and provide data-driven guidelines for training purposes [4]. However, additional sensors will increase the cost and may impair athletes' performance. Besides, one cannot ask his competitors to wear sensors in order to discover their strengths and weaknesses.

With the rapid development of computer vision technology, video analysis is more and more popular in helping improve athletes' expertise and training efficiency, and to prevent injuries [5]. This contactless technology enables coaches and athletes to train effectively, get rapid feedback, and minimize accidents. It can also help the coaches and athletes to analyze opponent teams' strength and weakness from past match videos, and design better strategies in future competitions [6].

The key steps for video analysis are player detection [7]. In this work, we develop a deep convolutional neural

network-based football video analysis algorithm. First, videos were converted to sequential images, which are then passed through five consecutive convolutional layers with batch normalization and leaky ReLU as the activation function in order to extract features with different levels of spatial resolution. A residual connection around the first three convolutional blocks followed by normalization was used to take into account all levels of feature maps and improve the detection accuracy. The upsampled feature maps were combined with feature maps of lower levels to obtain a player confidence map and a corresponding player bounding box.

2. Related Work

Traditional model player detection includes connected component analysis [8], shallow convolutional neural networks [9], histogram of orientated gradients and support vector machines (HOG-SVM) [10], and deformable part model (DPM) [11].

Figure 1 shows different situations in football player detection. Figure 1(a) is a typical image where players are separated with each other. Traditional models generally can detect players in this situation, while they can hardly detect adjacent players (Figures 1(b)–1(d)) correctly in a harder situation. Besides, HOG-SVM needs domain knowledge



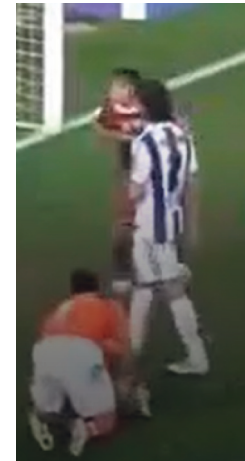
(a)



(b)



(c)



(d)

FIGURE 1: (a) A normal football match image for player detection. Difficult player detection tasks where players are nearby and (b) one player fell on the ground. (c) Plays' legs are overlapped. (d) One player is on his knees.

and more labor work in order to conduct background segmentation [10]. Non-maximum suppression restricts the performance of DPM when detecting close players [12]. Other approaches such as motion or pixel or template-based methods have various restrictions, such as the player should not stand still, cannot wear a white jersey, shorts, and socks [13].

With the rapid development of computer vision, neural networks are dominating object detection algorithms [14, 15]. YOLO object detection algorithms have evolved from version 1 to version 5 with improving capabilities and performances which beat traditional algorithms, such as DPM [16–20]. However, YOLO is a large neural network with millions of parameters to train, which restricted its application in real-time object detection on portable devices.

One main challenge in player detection is background distraction. Complicated backgrounds make small objects

harder to be detected. For example, if the background and players have similar colors, or players are too cluttered, it is hard to detect the players. Traditional methods use background subtraction approaches such as chromatic features [21], motion-based techniques [22], or median filtering [23] to preprocess the image frame.

Traditional deep neural networks suffer from vanishing gradients as the layers go deeper, leading to a worse performance than their shallow counterparts. He et al. reported using deep residual learning for image recognition [24], which overcome the vanishing gradient problem and enable training of very deep networks. The key part of the famous deep residual learning is the skip connection. The skip connection adds the output of the previous layer to the current layer and enables deep neural networks above 100 layers. The skip connection has been demonstrated to greatly enhance the model performance in image recognition.

In this work, we proposed a simple but efficient deep convolution neural network-based football player detection algorithm. It has two orders less parameters than YOLO and it can proceed an entire image in one pass. Besides, our model adopts a feature pyramid network design, which combines low level features with high level features. It helps differentiate an object with various sizes, and differentiate players with background clutter. A residual connection around the first three convolutional blocks followed by normalization was used to improve detection accuracy.

2.1. Model Architecture. Figure 2 presents the model architecture of our football player detector. This architecture is built on a feature pyramid network (FPN) with residual connections around the first 3 convolution neural network blocks. FPN is well-known for its high accuracy and high speed in object detection. It has been demonstrated to perform better than both ResNet and Faster R-CNN in many object detection tasks. The advantages of ResNet are that it can be trained easily even with a large number of layers and it can avoid the vanishing gradient problem by using residual connections. Here, we combine the advantages of ResNet and FPN to build a new architecture which inherits both models' advantages. Another benefit of using residual connection on FPN is that it enable our detector to make the final detection based on all levels of feature maps instead of the last level of feature maps. Lower levels of feature maps give more spatial location information than higher levels of feature maps. Also, we use denser grids to improve detecting accuracy when two players are closed to each other. The probability of presence of a player inside a grid cell was encoded into the player confidence map, and the coordinates of the player was encoded in the bounding box. To find the player position in the confidence map, we apply non-maximum suppression to the player confidence map and then filter out all the locations above a threshold. Furthermore, we combine a high-level feature and a large receptive field to improve the detection accuracy on players with different gestures, such as players on knees and players who fell on the ground.

Table 1 gives the detailed information of our model. Filters are applied to generate feature maps. Their functions are to help extract various features from an image, such as edges, horizontal and vertical lines, and curves. Max pooling was applied here after filters in order to select the most significant features in the patch and ignore the average features. Our experiments shows that max pooling gives much better results than average pooling and min pooling. From conv1 to conv5, the extracted features changes from broad features to very specific features.

3. Experiments

3.1. Dataset. We use a public dataset, ISSIA-CNR soccer and soccer player detection datasets, to train and evaluate our model.

We use random football match video clips obtained from Tiktok to test the generality of our model.

3.2. Loss Function. The loss function is comprised of two parts, player classification loss and bounding box loss.

$$L_p = - \sum_{(x,y) \in \text{positive}} \log c_{x,y} - \sum_{(x,y) \in \text{negative}} \log(1 - c_{x,y}), \quad (1)$$

where $c_{i,j}$ is the confidence score of the player at location (x, y) .

As shown in (1), the player loss is binary cross entropy. Positive means the player exists in that position, negative means that position (x, y) does not have any player.

$$L_{\text{bbox}} = \sum_{(x,y) \in \text{positive}} \text{smooth}_{L_1}(l(x, y) - g(x, y)), \quad (2)$$

where $l(x, y)$ represents a predicted boxing box in position (x, y) , and $g(x, y)$ represents the corresponding ground truth (labeled) bounding box.

For the bounding box loss, we use similar smooth L1 loss as described in the f-CNN paper [25].

$$L = \frac{1}{N} (L_p + \beta L_{\text{bbox}}), \quad (3)$$

where β is a hyperparameter that decides the weight of bounding box loss in the totally loss.

3.3. Model Training. Both datasets, ISSIA-CNR soccer and soccer player detection datasets, were used for training. We adopted a stratified train test split to reduce bias. In other words, 80% ISSA-CNR soccer and 80% soccer player detection datasets were selected randomly as the training set, 20% ISSA-CNR soccer and 20% soccer player detection datasets as the test dataset. Furthermore, we conducted cross-validation in order to reduce overfitting and improve the model performance. AdamW was used as the optimizer, and the learning rate scheduler was used to reduce the learning rate as training progresses.

3.4. Model Evaluation. We use the standard mean metric average precision (mAP) to evaluate the model. Intersection over union (IOU) of 0.5 was used as the threshold. Positive means IOU of the predicted bounding box and ground truth is higher than 0.5.

$$AP = \frac{1}{11} \sum \text{precision}(\text{recall}), \quad (4)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i.$$

3.5. Model Improvement. To improve the detection accuracy of our model on occlusion players, we have manually collected a large amount of occlusion football players from Tiktok and Youtube videos, manually labeled them, mixed them together with the ISSIA-CNR soccer and soccer player detection datasets to train our model.

To combat multiple spatial scales issues, we use anchor boxes to acquire various scale and aspect ratios of football

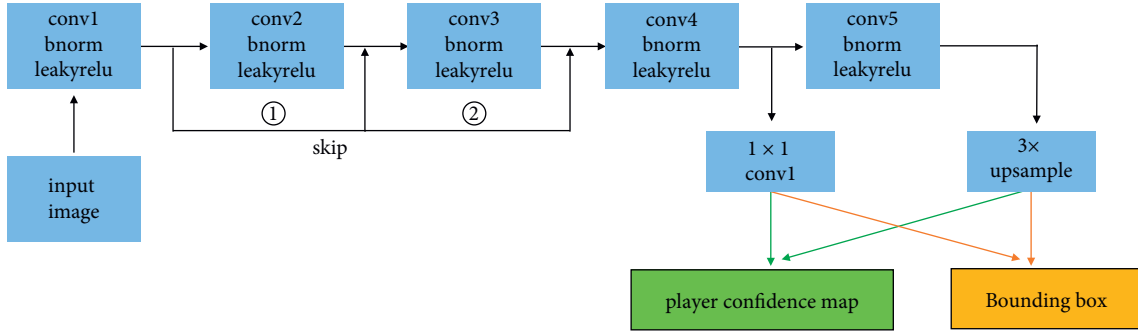


FIGURE 2: Model architecture of the football player detection system. First, an input image went through five consecutive convolutional layers with batch normalization and leaky ReLU as the activation function in order to extract features with different levels of spatial resolution. Then an upsampled feature map were combined with the feature level from the lower level to obtain the player confidence map and the corresponding player bounding box.

TABLE 1: Model details in terms of modules, layers and output dimensions.

Module	Layers	Output dimension
conv1	16 filters	[w/2, h/2, 16]
	Max pool 2 d (2 × 2) Residual connection	
conv2	32 filters	[w/4, h/4, 32]
	32 filters Max pool 2 d (2 × 2) Residual connection	
conv3	32 filters	[w/8, h/8, 32]
	32 filters Max pool 2 d (2 × 2)	
conv4	64 filters	[w/16, h/16, 64]
	64 filters Max pool 2 d (2 × 2)	
conv5	64 filters	[w/32, h/32, 32]
	64 filters Max pool 2 d (2 × 2)	
1 × 1 conv1	32 filters	[w/16, h/16, 32]
Player classifier	32 filters	[w/16, h/16, 1]
	2 filters	
	Sigmoid	
Bounding box	32 filters	[w/16, h/16, 4]
	4 filters	

players, combine feature maps from different convolutional layers, and adopt feature pyramid structures.

We make full use of residual connection to improve the speed of our model, while maintaining high detection accuracy. As the layers go deeper, accuracy will get increased, however, the training parameters grows drastically. The number of convolutional layers is optimized to be 5 in our work.

4. Results

4.1. Model Performance. Table 2 summarizes the evaluation results of our model on player detection using public datasets ISSIA-CNR and soccer player detection. Our model gives the highest AP score on both datasets (0.915 on the ISSIA-CNR dataset and 0.932 on the soccer player

detection dataset) with relatively less training parameters (238 k) and fast inference time (38 numbers of frames per second).

Specifically, for the ISSIA-CNR dataset, our model outperforms both the Faster R-CNN and FootAndBall model. The reason is that our architecture allows our model to capture feature maps with 5 levels of spatial resolution. The residual connections make feature maps from lower level conv layers flow to higher level conv layers easily and also helps the model to converge faster, which leads to shorter training time, fast frames processing, and higher prediction performance. Soccer player detection datasets are created from two different football matches with a wider range of pixels (20–250) than the ISSIA-CNR dataset (63–144 pixels). Our model beats the FootAndBall model, which is probably due to the abovementioned reasons that these residual connections allow all 5 levels of feature maps flow easily to the end before predicting the bounding box and the confidence score. In other words, the special architecture of our model makes it more robust and adapts our model to various football videos.

4.2. Effect of the Number of Convolutional Layers on the Model Performance. To further understand the functions of each block and optimize the model architecture, we extracted feature maps from different convolutional layers and feed them directly to the last 1 × 1 conv and 3x upsample blocks to generate the player confidence map and the bounding box. Then we evaluate the models using the same datasets (the soccer player detection dataset and the ISSIA-CNR dataset) and calculated the average precision. The results are shown in Figure 3.

Figure 3 shows that as convolution layers increase, the average precision of model prediction using both datasets increase. The average precision of the model from conv1 to conv5 on the ISSA-CNR dataset are 0.523, 0.718, 0.856, 0.893, and 0.902, respectively. The percentage enhancement of the average prediction of the model from conv2 to conv5 on the ISSA-CNR dataset compared with conv1 are 37%, 64%, 71%, and 72%, respectively. Similarly, the average precision of the model from conv1 to conv5 on the soccer player detection dataset are 0.557, 0.727, 0.847, 0.901, and

TABLE 2: Comparison of our model with literature models in terms of average precision (AP) of player detection, number of training parameters, and frames per second.

Model	ISSIA-CNR average precision	Soccer player detection average precision	Training parameters (k)	Frames per second	Reference
Faster R-CNN	0.874	0.928	25 600	8	[25]
FootAndBall	0.889	0.834	137	39	[26]
This work	0.915	0.932	238	38	NA

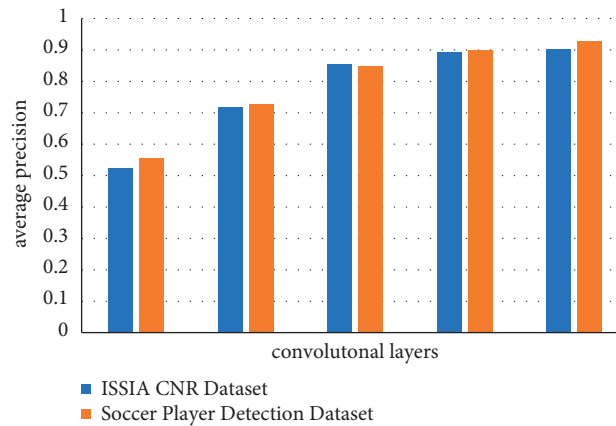


FIGURE 3: Model performance on two datasets using the feature map extracted from different convolutional layers. Conv1 corresponds to the feature map after the image pass conv1.

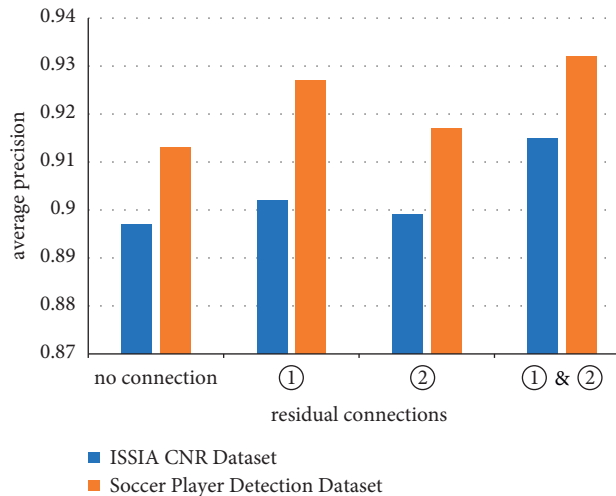


FIGURE 4: Effect of residual connections on the model performance.

0.927, respectively. The percentage enhancement of the average prediction of the model on the soccer player detection dataset from conv2 to conv5 compared with conv1 are 31%, 52%, 62%, and 66%, respectively. This clearly shows the importance of convolutional layers in object detection. However, as convolutional layers increase, adding more convolutional layers contribute less to the model performance, which is due to the fact that all important feature maps have been extracted. It also indicates that 4–5 layers of convolutional layers is good enough for player object detection tasks.

4.3. Effect of the Residual Connection on the Model Performance. As shown in Figure 4, when there is no residual connection, the model performance is the worst. In fact, the training time is also longer in this case. Adding residual connection 1 increased the model performance much more than adding residual connection 2, which is probably because the feature map after conv3 contains much information from both conv2 and conv1 (comes from residual connection 1), and thus it helps improve the model performance from an average precision of 0.897 to 0.902 (0.55%) for the ISSIA_CNR dataset and from an average



FIGURE 5: Model performance on (a) a normal football match image for player detection, and on difficult player detection tasks where players are nearby and (b) one player fell on the ground. (c) Plays' legs are overlapped. (d) One player is on his knees.

precision of 0.913 to 0.927 (1.53%) for the soccer player detection dataset. However, if we only add residual connection 2, the enhancement of the model performance is not so much, which maybe because the spatial resolution of feature maps between conv2 and conv3 have less difference than these between conv1 and conv2. Thus, residual connection 2 did not contribute too much information of feature maps with different spatial resolution. When residual 1 and residual 2 are all added, we obtained the best performance, an average precision of 0.915 for the ISSIA_CNR dataset and 0.932 for the soccer player detection dataset.

To further prove the generality of our model, we collect random football match clips from Tiktok and feed it directly into our model for player detection. Figure 5 shows the model performance on these unseen images from Tiktok. In Figure 5(a), our model can detect all the football players with a confidence score close to 1, which outperforms traditional models. In Figure 5(b), even though the image patch is blurry

and one player falls to the ground, our detector still successfully detects both players with a confidence score of 0.78. This is because we use denser grids compared to other popular models, such as YOLO. We scale down the input image size by a factor of 16. In Figure 5(c), two players' bodies are partially overlapped with each other, our model detects both players based on only a part of their bodies and gives a confidence score of 0.88 and 0.85, respectively. In Figure 5(d), one player is on his knees, but our model successfully detect him with a confidence score of 0.84. The high performance of our model on a random football match video clip and on difficult tasks assures the generality and effectiveness of our algorithm.

5. Conclusions

In conclusion, we have proposed an efficient deep convolutional neural network-based method to automatically detect football players from video matches directly. Our

network was built on the pyramid network with residual connections, with the advantages of single pass fast processing, high robustness, adaptive to all size of images, and suitable for nearly any match videos. Our results shows that it can perform well even on random football match videos obtained from Tiktok, indicating the wide applications of this algorithm. Moreover, our player detection algorithm is faster than the state-of-the-art R-CNN object detector and can be used for real-time football player detection. In the future, we plan to combine transformers with our current deep convolution neural network to not only detect football players, but also predict each player's next action and state.

Data Availability

The dataset is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Fundamental Research Funds in Heilongjiang Provincial Universities of China under Grant no. 135309418.

References

- [1] C. Zhu, R. Shao, X. Zhang, S. Gao, and B. Li, "Application of virtual reality based on computer vision in sports posture correction," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 3719971, 1–15 pages, 2022.
- [2] L. Zhu, "Computer vision-driven evaluation system for assisted decision-making in sports training," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1865538, 1–7 pages, 2021.
- [3] M. Burić, M. Pobar, and M. Ivašić-Kos, "Object detection in sports videos," in *Proceedings of the 2018 41st International Convention on Information and Communication Technology*, pp. 1034–1039, Opatija, Croatia, 21–25 May 2018.
- [4] P. Salvo, A. Pingitore, A. Barbini, and F. Di Francesco, "A wearable sweat rate sensor to monitor the athletes' performance during training," *Science & Sports*, vol. 33, no. 2, pp. e51–e58, 2018.
- [5] J. T. Johnston, B. R. Mandelbaum, D. Schub et al., "Video analysis of anterior cruciate ligament tears in professional American football athletes," *The American Journal of Sports Medicine*, vol. 46, no. 4, pp. 862–868, 2018.
- [6] I. McKeown, K. Taylor-McKeown, C. Woods, and N. Ball, "Athletic ability assessment: a movement assessment protocol for athletes," *International journal of sports physical therapy*, vol. 9, no. 7, pp. 862–873, 2014.
- [7] J. Liu, X. Tong, W. Li, T. Wang, Y. Zhang, and H. Wang, "Automatic player detection, labeling and tracking in broadcast soccer video," *Pattern Recognition Letters*, vol. 30, no. 2, pp. 103–113, 2009.
- [8] R. G. Abbott and L. R. Williams, "Multiple target tracking with lazy background subtraction and connected components analysis," *Machine Vision and Applications*, vol. 20, no. 2, pp. 93–101, 2009.
- [9] A. Lehuger, S. Duffner, and C. Garcia, "A robust method for automatic player detection in sport videos," *Orange Labs*, vol. 4, 2007.
- [10] S. Maćkowiak, M. Kurc, J. Konieczny, and P. Maćkowiak, "A complex system for football player detection in broadcasted video," in *Proceedings of the ICSES 2010 International Conference on Signals and Electronic Circuits*, pp. 119–122, Gliwice, Poland, 07–10 September 2010.
- [11] D. Zhang, "Vehicle target detection methods based on color fusion deformable part model," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 94, 2018.
- [12] W. Wei-Lwun Lu, J. A. TingTing, J. J. Little, and K. P. Murphy, "Learning to track and identify players from broadcast sports videos," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1704–1716, 2013.
- [13] M. Manafifard, H. Ebadi, and H. Abrishami Moghaddam, "A survey on player tracking in soccer videos," *Computer Vision and Image Understanding*, vol. 159, pp. 19–46, 2017.
- [14] Z. Cai, Q. Fan, R. S. Feris, and N. Vasconcelos, "A unified multi-scale deep convolutional neural network for fast object detection," in *Proceedings of the 14th European Conference*, pp. 354–370, Amsterdam The Netherlands, October 11–14, 2016.
- [15] A. Dhillon and G. K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," *Progress in Artificial Intelligence*, vol. 9, no. 2, pp. 85–112, 2020.
- [16] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788, Las Vegas, NV, USA, 27–30 June 2016.
- [17] Z. Hui-juan, Z. Qiang, L. Yu, W. Xu-yang, and L. Ying, "Face detection method based on YOLO2 for subway passenger flow into station," *Computer and Modernization*, vol. 76, 2019.
- [18] T. Li, Y. Ma, and T. Endoh, "A systematic study of tiny YOLO3 inference: toward compact brainware processor with less memory and logic gate," *IEEE Access*, vol. 8, pp. 142931–142955, 2020.
- [19] D. Wu, S. Lv, M. Jiang, and H. Song, "Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments," *Computers and Electronics in Agriculture*, vol. 178, Article ID 105742, 2020.
- [20] G. Yang, W. Feng, J. Jin et al., "Face mask recognition system with YOLOV5 based on image recognition," in *Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, 2020, pp. 1398–1404, Chengdu, China, 11–14 December 2020.
- [21] H. S. Y. YoonYoon, Y. I. J. B. BaeBae, and Y. k. Y. YangYang, "A soccer image sequence mosaicking and analysis method using line and advertisement board detection," *ETRI Journal*, vol. 24, no. 6, pp. 443–454, 2002.
- [22] P. L. Mazzeo, M. Leo, P. Spagnolo, and M. Nitti, "Soccer ball detection by comparing different feature extraction

- methodologies,” *Advances in Artificial Intelligence*, vol. 2012, Article ID 512159, 1–12 pages, 2012.
- [23] P. R. Kamble, A. G. Keskar, and K. M. Bhurchandi, “A deep learning ball tracking system in soccer videos,” *Opto-Electronics Review*, vol. 27, no. 1, pp. 58–69, 2019.
- [24] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, July 2016.
- [25] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-CNN,” in *Proceedings of the IEEE International Conference on Computer Vision*, Honolulu, HI, USA, June 2017.
- [26] J. Komorowski, G. Kurzejamski, and G. Sarwas, “FootAndBall: Integrated Player and ball Detector,” 2019, <https://arxiv.org/abs/1912.05445>.

Research Article

Recommendation of Business Models for Agriculture-Related Platforms Based on Deep Learning

Yufei Zhou and Sha Hua 

School of Business, Hunan Agricultural University, Changsha 410128, China

Correspondence should be addressed to Sha Hua; hs987122141@stu.hunau.edu.cn

Received 3 June 2022; Accepted 28 June 2022; Published 11 July 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yufei Zhou and Sha Hua. 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.

Agriculture is a basic and pillar industry. With the integration and development of Internet+, platform economy, and various industries, the business model of agriculture-related platforms is also constantly innovating. In this context, it is necessary to recommend suitable business models for different types of agriculture-related platforms. Based on the characteristics of agriculture-related platforms and various business models, this paper proposes a business model recommendation algorithm based on radial basis function neural network (RBFNN). This method trains the RBFNN model with the goal of maximizing the correlation between agricultural-related platforms and business models. In the application stage, for a specific agriculture-related platform, after inputting its characteristic parameters, a suitable business model can be recommended. In the experiment, the proposed method is tested and verified with relevant data, and the results show the effectiveness of the method.

1. Introduction

As the consumption revolution triggered by the digital network and the platform economy continues to impact traditional industries, the agriculture-related platforms that combine Internet technology with agriculture have also developed in full swing in the past years [1, 2]. Chinese agriculture is a combination of tradition and modernity. Generally speaking, it is still the type of small farmers in East Asia. It can comprehensively utilize extremely limited resources to the extreme. However, the decentralized agricultural structure is not conducive to the introduction of modern production factors. Also, it is difficult to achieve standardization and large-scale production. How to transform Chinese agriculture from a consumption-based production mode that is overly dependent on resources to a high-quality sustainable green modern production mode is a strategic issue, which needs to be solved urgently in current theoretical research and corporate practice. With the mobile Internet, the large e-commerce companies seriously affected the development of small farmers. However, community

group buying cannot fully meet the needs of consumers, and there is a market gap in the production and marketing of agricultural products. However, the development trend of “Internet + agriculture” and agriculture-related platforms is unstoppable. Although agriculture-related platforms are strongly supported by policies, there is still a lack of relevant research on their specific business models.

Early recommendation systems mainly used collaborative filtering (CF) algorithms [3–5]. Subsequently, the recommendation methods based on machine learning algorithms such as support vector machine (SVM) have also been studied and applied [6–8]. In recent years, the deep learning techniques have provided powerful tools for data analysis, and typical models include convolutional neural network (CNN), recurrent neural network (RNN), long short-term memory (LSTM), and so on [9–13]. Relevant applications show that the deep learning methods have superior performance in the field of recommendation algorithms. On the basis of related research, this paper proposes a business model recommendation method for agriculture-related platforms based on radial basis function

neural network (RBFNN) [14–16]. The proposed method trains the designed RBFNN with the goal of optimizing the correlation between agricultural-related platforms and business models. On this basis, for the input agriculture-related platform, the best business model can be recommended according to its relevant parameters. The experiments are carried out based on the relevant datasets of agricultural platforms. The results show that the recommendation results of the method in this paper are reliable and have performance advantages compared with several existing methods.

2. Background

The business model is the main body and carrier of innovation in the mobile Internet era. It is a complex phenomenon with multiple factors coexisting and intertwined in multiple dimensions, which is a transaction activity to realize a specific value proposition and an important medium for the market and technology. It is widely used in management practice and research and has become a hot research topic in various fields such as organization, strategy, and entrepreneurship. The study of business models begins with the attention to the formation and design of business models. Also, it has gone through many iterations of interpretations by different scholars on the definitions, connotations, and elements of related concepts. First of all, from the definition point of view, the authors in [17] first proposed that a business model is an organic system composed of services, products, and information, which was a unique combination of tangible and intangible assets of an enterprise. Subsequently, the authors in [18] elaborated from the perspective of function, arguing that business model was an important carrier of the ability of enterprises to continuously obtain profits and transformation and upgrading, as well as the key factor to generate enterprise performance. Saebi and Foss [19], through the study of e-commerce model, proposed that a business model was a system of interdependent operational activities between focal enterprises and stakeholders.

In terms of business model innovation, since 2012, some scholars have combined innovation and business model themes and advanced it in depth, which has gradually become the mainstream direction of business model research and has good research and development prospects [20–23]. The innovation of business models is not only an activity of transaction innovation but also a process of institutional innovation and reform in which entrepreneurs and stakeholders interact to create, give, and construct meaning. It is a high-level enterprise innovation behavior including the internal and external factors. It is very different from product innovation, technological innovation, system innovation, and conceptual innovation in the traditional sense. It is a dynamic process of interaction between enterprises and the internal and external environment of the market. Also, there is a changing path that plays a key role in business model innovation. Magretta [20] proposed a method to reconstruct or improve the interaction between the internal components of the business model by redesigning any of the

basic elements of the nine elements in its business canvas, such as changing the value proposition of distribution channels and redesigning resource elements. Teece [21] believed that through the innovation or reconstruction of internal elements such as customer segmentation standards, transaction rules, and the relationship between the demander and the supplier, the transformation of the entire transaction system and the innovation of business models can be achieved. Casadesus and Zhu [22] believed that the innovation of internal value drivers and external driving factors can promote the overall innovation of business models.

The research perspective of existing business models mainly starts from the perspective of transaction activities or value creation. The path of business model innovation generally includes element innovation, value proposition innovation, operation model innovation, and business system structure innovation. The authors believe that a business model is the connection transaction method and interest structure between enterprises and stakeholders, which focuses on the value co-creation mechanism of platform business model. The dynamic perspective is mostly based on the value chain, which is difficult to apply to the complex value network created by the platform enterprises through the continuous integration of cross-border resources around the core platform. The static perspective of business model innovation emphasizes the identification and re-emphasis within the business model. It is necessary to construct key elements and core components to realize business model innovation. Also, the research results are too theoretical and not conducive to guiding practice. This research is dedicated to deconstructing and analyzing business model innovation from a systematic perspective, combining static research and dynamic research on business model innovation paths. Therefore, the path of business model innovation is first logically deduced from the perspective of the combination of transaction activities and value creation. By extracting external value drivers and internal attributes, it explains the mechanism and specific change process of the innovation of the platform-based business model of agricultural cooperatives. The core operation logic is that enterprises meet the potential value needs of customers through value proposition innovation, form value maintenance, and ultimately realize the value creation of the entire value network. The constituent units in its framework cover the realization path of business model innovation. Through the design of transaction structure and value creation logic, new value propositions are formed to develop business opportunities and broaden value sources. The platform business is realized through the unique business system structure of value network. Co-creation and sharing of the value of each participant in the ecosystem can increase the added value of consumers' cognitive attachment to products or services. Through the rational allocation of resources, a network effect, scale effect, and monopoly value can be formed. Value co-creation makes economic activities and value transmission change from the one-way path of the traditional value chain to the relationship of value network, which accelerates the efficiency of information

dissemination, reduces transaction costs, differentiates products and advantages, and effectively reduces the malicious competition of homogeneous products.

3. Method Description

3.1. Similarity Model. The business model recommendation problem model involved in this paper includes two entity models: agriculture-related platform and business model. The recommendation problem is actually based on the characteristics of the multiple dimensions of the agriculture-related platform and the business model, to find the minimum difference in the characteristics between the two. Then, a proper way is found to perform accurate matching and to recommend the best business model. Considering the characteristics of agriculture-related platforms and business models, this paper defines the main characteristics of agriculture-related platforms and business models as listed in Table 1.

According to the platform and business model features in the above table, they are first digitally encoded and then we compare the features to find the difference value. Afterwards, the four corresponding feature difference functions $S_i (i = 1, 2, \dots, 4)$ can be obtained, so the recommended model objective function is formed as follows:

$$\min S = \sum_{i=1}^4 \omega_i S_i, \quad (1)$$

where ω_i represents the weight of different difference functions to the overall recommendation model. According to the recommendation model objective function, for each platform, all available resource features and platform feature differences are sorted in ascending order, and the top-ranked business model is selected to generate a recommendation sequence.

3.2. RBFNN. Aiming at the above optimization problem, this paper adopts RBFNN as the basic method to solve it. RBFNN includes a three-layer structure of input layer, hidden layer, and output layer. The relationship between the input layer and the hidden layer is nonlinear, and the relationship between the hidden layer and the output layer is linear. RBFNN uses the radial basis function as the activation function of the neurons in the hidden layer to map an appropriate amount of input to the latent space. Therefore, as long as the center point of the radial basis function is determined, the mapping relationship between the input layer and the hidden layer will also be determined. The hidden layer and the output layer are connected by weights, where the weights are network adjustable parameters. The structure of the RBFNN is shown in Figure 1.

Suppose that the input vector is $X = [x_1, x_2, \dots, x_n]^T$, n is the number of input layer units, the output vector is $Y = [y_1, y_2, \dots, y_q]^T$, q is the number of output layer units, and the number of hidden layer units is p . Here the Gaussian function is used as the kernel function, so the output value of the hidden layer neuron is as follows:

$$h_j = \exp\left(-\left\|\frac{X - C_j}{D_j}\right\|^2\right), \quad j = 1, 2, \dots, p, \quad (2)$$

where $C_j = [c_{j1}, c_{j2}, \dots, c_{jn}]^T$ is the center vector of the hidden layer neuron j and D_j is the width vector of the hidden layer neuron j , which is related to the action range of the hidden layer neuron on the input; the smaller D_j , the narrower the activation function of the neuron j ; the other neurons have less influence on neuron j .

It can be seen from the structure of RBFNN that the input layer and the hidden layer are directly connected, and the hidden layer and the output layer are connected through the weight matrix, which can calculate the initial value of the central parameter, the initial value of the weight, and the width vector. The training process is to solve the network center vector, width vector, and weight matrix of the model.

Figure 2 shows the basic flow of RBFNN training. First, the neural network is initialized. Then, the output results and the root mean square error are calculated with the expected output. If the error meets the termination condition, the training is terminated. Otherwise, the gradient descent method is used to continue to adjust C_j , D_j , and W_k ; then, perform cyclic calculation and judge whether the termination condition is met.

3.3. Model Realization. According to the above process, the basic process of developing the agricultural platform business model in this paper is described as follows. First, the learning record samples are input. The characteristics of the agricultural platform and business model are encoded and the characteristic variables are input into the RBFNN. The core parameters of the RBFNN are optimized through the algorithms in 3.1 to obtain a stable RBFNN recommendation model. Finally, through multiple training, a stable business model recommendation model for agriculture-related platforms is obtained. The specific condition for iterative stop is whether the feature difference meets the requirements, that is, whether the feature difference reaches the minimum value. According to the trained recommendation model, when a new agriculture-related platform is input, its features can be directly input into the recommendation model, so as to obtain the best business model recommendation result.

4. Experiment and Analysis

4.1. Dataset. In order to verify the performance of the method in this paper in the recommendation of agricultural platform business model, this section mainly conducts experimental verification. The experimental samples include 4,387 items. Among them, 10 typical business models are included. In the experiment, several types of classic recommendation models in the existing literature are selected as comparison methods, including CF, SVM, and LSTM.

In order to quantitatively evaluate and compare the performance of various methods, the experiments use evaluation indicators widely used in existing literature,

TABLE 1: Characteristics of agriculture-related platforms and business models.

	Feature	Description
Platform	Content	Specific agricultural areas
	Scale	Funds, personnel, etc.
	Region	Location
	Qualification	Various business qualifications, etc.
Business model	Content category	Agricultural direction involved
	Difficulty level	Requirements for the qualification and scale of the platform
	Popularity	Frequency of use on other platforms
	Region	Applicable regions

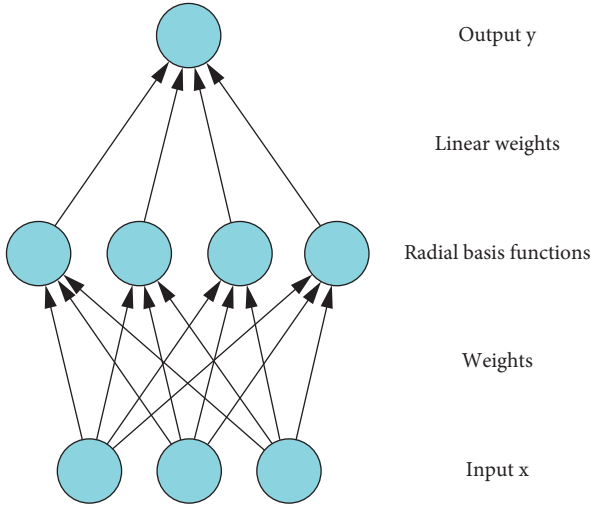


FIGURE 1: Basic structure of RBFNN.

including precision (P), recall (R), and F_1 -score (F_1). The definitions of three indexes are as follows:

$$\begin{aligned}
 P &= \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|}, \\
 R &= \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|}, \\
 F_1 &= \frac{2 \times P \times R}{P + R},
 \end{aligned} \quad (3)$$

where $R(u)$ represents the business model selected according to the agricultural platform in the training dataset and $T(u)$ represents the user's business model recommendation list on the test dataset.

4.2. Result and Analysis. Based on the above database, the proposed method and the comparison methods are tested. Also, the statistical results are shown in Table 2. From the results, under the three evaluation indicators, the method in this paper has achieved the best performance, showing its effectiveness. Among the 3 types of comparison methods, the deep learning-based LSTM method has the best performance, reflecting the reliability of the deep learning method in data analysis and prediction. Compared with

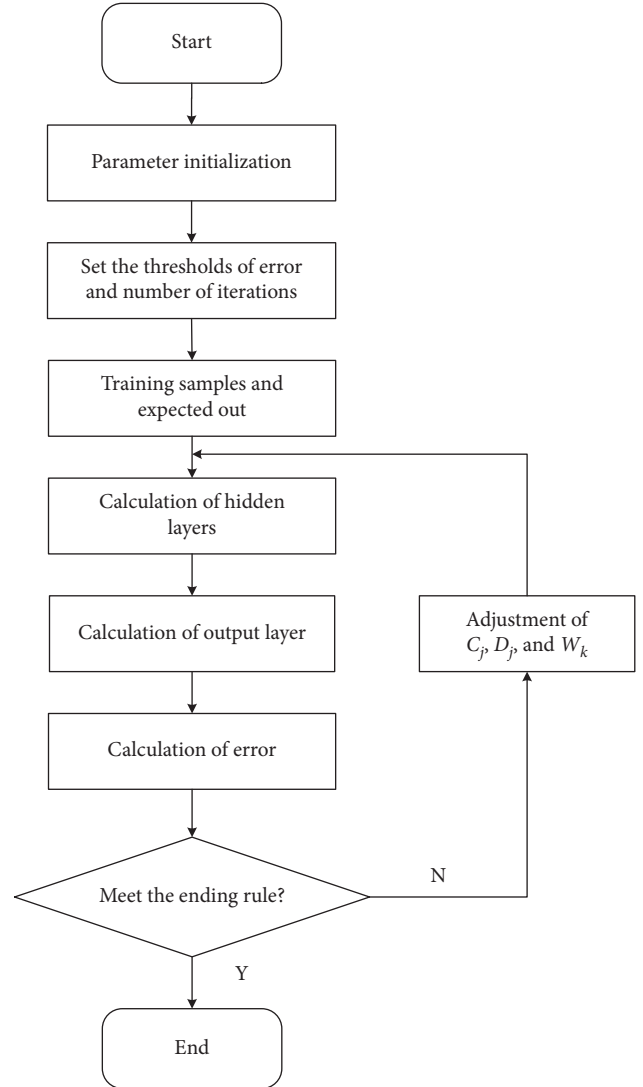


FIGURE 2: Basic flowchart of RBFNN training.

LSTM, the method in this paper improves the three indicators of P , R , and F_1 by 2.3%, 2.1%, and 1.7%, respectively, showing the important role of RBFNN in the recommendation algorithm. In conclusion, the method in this paper can be effectively applied to the business model recommendation of agriculture-related platforms and can achieve good results.

TABLE 2: Performance comparison of different methods.

Method	P	R	F_1
Proposed	0.824	0.873	0.851
CF	0.732	0.758	0.742
SVM	0.727	0.749	0.735
LSTM	0.801	0.852	0.834

5. Conclusion

Aiming at the business model recommendation problem of agriculture-related platforms, this paper proposes a recommendation model based on RBFNN. First, the correlation function between agriculture-related platforms and different business models is established, so as to obtain the objective function that needs to be optimized. Then, with the goal of maximizing the objective function, the designed RBFNN is trained to build a recommendation model. For the input agricultural platform, the best business model can be recommended according to its own characteristics. The proposed method is verified by experiments and compared with several existing recommendation algorithms. The results show the effectiveness of the method. By using relevant systems in the development of agriculture-related platforms, it can contribute to the healthy and stable development of agriculture-related platforms.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] N. Nawaz Malik, W. Alosaimi, M. Irfan Uddin, B. Alouffi, and H. Alyami, "Wireless sensor network applications in healthcare and precision agriculture," *Journal of Healthcare Engineering*, vol. 2020, Article ID 8836613, 9 pages, 2020.
- [2] M. Grossi, A. Berardinelli, E. Sazonov, W. Beccaro, and O. Martin, "Sensors and embedded systems in agriculture and food analysis," *Journal of Sensors*, vol. 2019, Article ID 6808674, 2 pages, 2019.
- [3] W. Yang, "Personalized intelligent recommendation algorithm design for book services based on deep learning," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 9203665, 8 pages, 2022.
- [4] Z. H. Wang and Z. DeHou, "Research on book recommendation algorithm based on collaborative filtering and interest degree," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 7036357, 7 pages, 2021.
- [5] S. Zhang, L. N. Yao, A. X. Sun, and Y. Tay, "Deep learning based recommender system: a survey and new perspectives," *ACM Computing Surveys*, vol. 52, no. 1, p. 5, 2019.
- [6] B. Yang, Y. Lei, and J. M. Liu, "Social collaborative filtering by trust," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, pp. 1633–1647, 2017.
- [7] L. H. Jonathan, A. K. Joseph, and R. John, "Explaining collaborative filtering recommendations," in *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, pp. 241–250, Hangzhou, March 2001.
- [8] F. Strub, R. Gaudel, and J. Mary, "Hybrid recommender system based on autoencoders," in *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, pp. 11–16, ACM, Boston, MA, USA, September 2016.
- [9] J. Li and Z. Ye, "Course recommendations in online education based on collaborative filtering recommendation algorithm," *Complexity*, vol. 2020, Article ID 6619249, 2020.
- [10] X. Liu, "Music trend prediction based on improved LSTM and random forest algorithm," *Journal of Sensors*, vol. 2022, Article ID 6450469, 10 pages, 2022.
- [11] Z. H. Cui, X. H. Xu, F. Xue et al., "Personalized recommendation system based on collaborative filtering for IoT scenarios," *IEEE Transactions on Services Computing*, vol. 13, no. 4, pp. 685–695, 2020.
- [12] S. Sedhain, A. K. Menon, S. Sanner, and L. Xie, "Autorec: autoencoders meet collaborative filtering," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 111–112, ACM, Florence Italy, May 2015.
- [13] J. Zhang, B. W. Hao, B. Chen, C. Li, H. Chen, and J. Sun, "Hierarchical reinforcement learning for course recommendation in MOOCs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 435–442, Honolulu HI USA, February 2019.
- [14] Y. Tian, K. Zhang, J. Li, X. Lin, and B. Yang, "LSTM-Based traffic flow prediction with missing data," *Neurocomputing*, vol. 318, no. 27, pp. 297–305, 2018.
- [15] N. J. Foss and T. Saebi, "Fifteen years of research on business model innovation," *Journal of Management*, vol. 43, no. 1, pp. 200–227, 2017.
- [16] S. M. Shafer, H. J. Smith, and J. C. Linder, "The power of business models," *Business Horizons*, vol. 48, no. 3, pp. 199–207, 2005.
- [17] C. M. DaSilva and P. Trkman, "Business model: what it is and what it is not," *Long Range Planning*, vol. 47, no. 6, pp. 379–389, 2014.
- [18] R. Casadesus-Masanell and J. E. Ricart, "How to design a winning business model," *Harvard Business Review*, vol. 89, no. 1-2, pp. 100–107, 2011.
- [19] T. Saebi and N. J. Foss, "Business models for open innovation: matching heterogeneous open innovation strategies with business model dimensions," *European Management Journal*, vol. 33, no. 3, pp. 201–213, 2015.
- [20] J. Magretta, "Why business models matter," *Harvard Business Review*, vol. 80, no. 5, pp. 86–92, 2002.
- [21] D. J. Teece, "Business models, business strategy and innovation," *Long Range Planning*, vol. 43, no. 2-3, pp. 172–194, 2009.
- [22] M. R. Casadesus and F. Zhu, "Business model innovation and competitive imitation: the case of sponsor-based business models," *Strategic Management Journal*, vol. 34, no. 4, pp. 464–482, 2013.
- [23] C. Zott and R. Amit, "Business model design: an activity system perspective," *Long Range Planning*, vol. 43, no. 2-3, pp. 216–226, 2010.

Research Article

Music Score Recognition Method Based on Deep Learning

Qin Lin 

Art College of Guizhou University of Finance and Economics, Guiyang 550001, Guizhou, China

Correspondence should be addressed to Qin Lin; linqin@mail.gufe.edu.cn

Received 18 February 2022; Accepted 22 March 2022; Published 7 July 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Qin Lin. 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 recent years, the recommendation application of artificial intelligence and deep music has gradually become a research hotspot. As a complex machine learning algorithm, deep learning can extract features with value laws through training samples. The rise of deep learning network will promote the development of artificial intelligence and also provide a new idea for music score recognition. In this paper, the improved deep learning algorithm is applied to the research of music score recognition. Based on the traditional neural network, the attention weight value improved convolutional neural network (CNN) and high execution efficiency deep belief network (DBN) are introduced to realize the feature extraction and intelligent recognition of music score. Taking the feature vector set extracted by CNN-DBN as input set, a feature learning algorithm based on CNN&DBN was established to extract music score. Experiments show that the proposed model in a variety of different types of polyphony music recognition showed more accurate recognition and good performance; the recognition rate of the improved algorithm applied to the soundtrack identification is as high as 98.4%, which is significantly better than those of other classic algorithms, proving that CNN&DBN can achieve better effect in music information retrieval. It provides data support for constructing knowledge graph in music field and indicates that deep learning has great research value in music retrieval field.

1. Introduction

With the rapid development of the Internet and information society, intelligent, globalized, and nongroup information features widely exist in our life. The information age makes us transfer resources in a more convenient way. In the era of big data, with the development of data mining technology, the technology of mining and refining massive data and converting it into valuable information has gradually become popular and mature. People get a lot of video, audio, and text information through the computer network and get spiritual enjoyment from it, among which music is very important for human beings [1]. With the development of the Internet era, people require the retrieval of music information to be more rapid, accurate, convenient, and efficient. However, music has many characteristics of its own, such as tone, melody, timbre, and tone strength, but the characteristics of music itself are difficult to be embodied, unlike traditional texts that can be described in an intuitive way. With the development of digital media economy, music information retrieval and

recognition has become a hot research field for many scholars. With the emergence of deep learning, music classification technology has entered a new development period [2, 3]. Deep learning in areas such as image processing and speech recognition has been more widely used than the traditional machine learning method, and the scholars have also started to use deep learning technology research in the field of music information retrieval related issues. Not only is the label of musical instrument of great significance to the classification of music type, but also it can be used to predict the emotion and music scene contained in music, so the identification and classification of musical instrument also play an important role in the field of music information retrieval. If we know the instrument that plays a particular piece of music, then we can optimize the automatic classification of music according to the characteristics of the instrument used [4]. Because the research on musical instrument recognition and classification can help people to study the recognition and classification of other music information retrieval fields, the research on musical instrument recognition and

classification has attracted more and more attention of scholars recently.

The proposal and application of the concept of deep learning is the biggest breakthrough in the 21st century. In 2011, Microsoft used deep neural network (DNN) for the first time, which has achieved remarkable results in speech recognition tasks. DNN is attracting more and more attention in speech recognition. Deep learning (DL) is the development of artificial neural network, which has a good application effect in the fields of speech and image and is currently considered a research hotspot. With the development of the times, the amount of network multimedia audio is increasing day by day. How to effectively retrieve the music information that people need has become an urgent problem to be dealt with by music information retrieval technology. Before the emergence of deep learning technology, the research on music classification mainly focused on the selection of feature sets and classifiers [5]. With the emergence of open music datasets, there are more and more research works in the field of music classification. However, most of the research works are still limited to the methods that take the features of the bottom of audio as feature sets and traditional machine learning algorithms as classifiers. The correlation between audio features and music categories is elusive, and the effect of classification using these underlying features is not stable, which largely depends on the selected feature set, and it is difficult to deal with large-scale music data using traditional classification algorithms. The mentioned problems limit the development of music classification research. With the rise of deep learning, classification methods based on deep neural network begin to emerge [6, 7]. Musical instrument recognition is an important field of audio retrieval, which involves not only the acoustic properties of sound source but also the perception psychology of human ear.

Based on previous studies, on the basis of considering the deep learning strong ability of feature extraction in the field of image recognition, this paper explores the innovative deep learning algorithms in the research and application of the background music classification recognition; on the basis of the classical neural network, the eigenvalue vector was introduced; first of all is the preprocessing of sample datasets combined with characteristic parameters change, followed by strengthening the characteristic value; then, the classification model of music score after deep supervised learning is established, and the model application comparison test is carried out on the test samples. The experiment proves that the recognition rate of the improved algorithm in music recognition is as high as 98.4%, which is obviously better than those of other classical algorithms. The main innovation of this paper is that the proposed model shows more accurate recognition ability and good execution performance in recognition of multiple types of polyphony music.

2. Basic Theory of Deep Learning

Deep learning itself is a kind of artificial neural network, on which many improvements have been made to deepen the depth and improve the complexity of the network. It can be

said that deep learning is a series of related technologies generated for better application of deeper and more complex networks [8]. Neurons in deep learning models mimic the way neurons work in the brain, trying to mimic the way stimuli communicate with neurons in the brain. Since the development of deep learning, many classical network structures have been born, and they have achieved achievements that traditional machine learning could not achieve. Convolutional neural network adopts the structure pattern of human neural tissue, sets up the convolution layer, and provides the associated subsampling layer for it. It uses certain rules to connect the upper and lower layers and the adjacent neural tissues to establish a forced and local association relationship in each layer. These units have a continuous receiving domain in space.

Convolutional neural network has the characteristics of reducing network parameters and learning local features. Compared with traditional fully connected neural network, the biggest difference is that only a part of nodes are connected between adjacent layers, so as to avoid the problem of too many parameters in the fully connected layer. By reducing the parameters, the convolutional neural network can calculate faster and avoid the overfitting problem [9]. The typical structure of CNN is shown in Figure 1. Figure 1 is reproduced from the work of Chen and Tang (2021) [10] (under the Creative Commons (Attribution License/public domain)).

2.1. Input Layer. Consider the input of a channel number of data F_0 ; the standard image processing $F_0 = 3$, exactly corresponding to the three color channels of the computer: red, green, and blue [11]. Set the input after processing as \tilde{X} , and the specific calculation process can be expressed as

$$\tilde{X}_{f,j,k}^{(i)} = \frac{X_{f,j,k}^{(i)} - \hat{X}_{f,j,k}^{(i)}}{\sigma_{f,j,k}}, \quad (1)$$

where $\hat{X}_{f,j,k}^{(i)} = 1/l \sum_{i=0}^{l-1} X_{f,j,k}^{(i)}$ and $\sigma_{f,j,k} = \sqrt{1/l \sum_{i=0}^{l-1} (X_{f,j,k}^{(i)} - \hat{X}_{f,j,k}^{(i)})^2}$.

2.2. Padding. The completion operation of size is P , 1 is added to the beginning and end of each row and column of the feature graph [12]. The padding of the feature map is shown in Figure 2; the red part represents the part of completion, where $P = 1$.

2.3. Convolution. The convolution layer is the core network layer in the convolutional neural network. It mainly carries out the convolution operation, extracts local features based on the spatial local correlation of input data, and forms the overall features by connecting these local features. CNN is named from the unique convolution operation in the model, which is also the most important part of the model. Figure 3 illustrates the convolution operation diagram. Figure 3 is reproduced from the work of Chen and Tang (2021) [10] (under the Creative Commons (Attribution License/public domain)). As can be seen from the figure, neurons at each

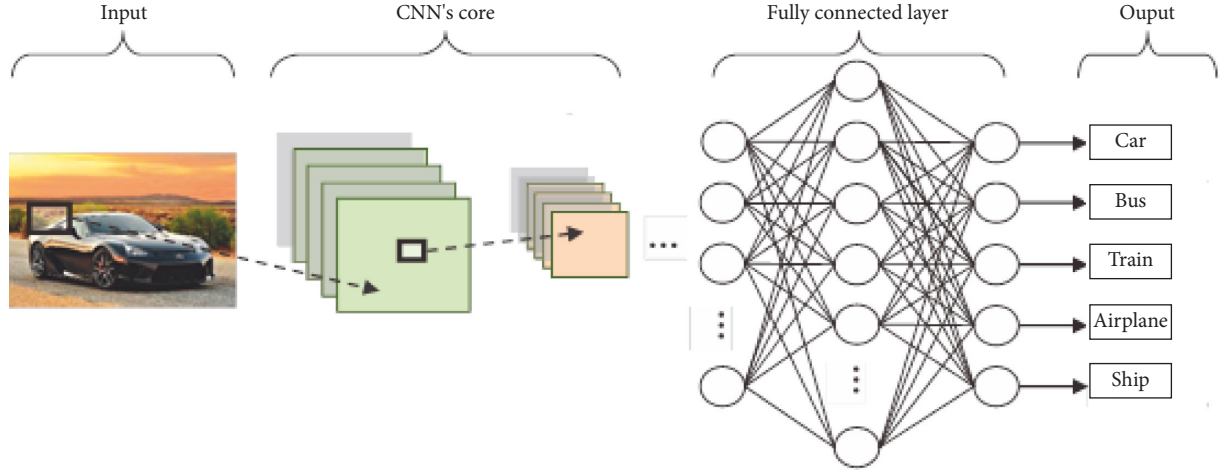


FIGURE 1: Typical structure of CNN [10].

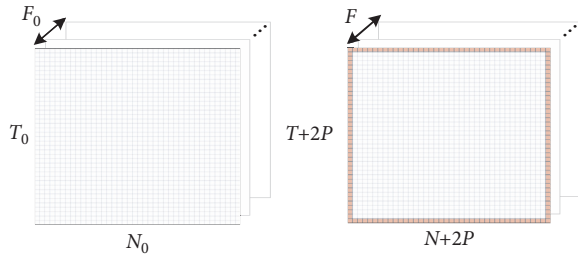


FIGURE 2: Padding of the feature map.

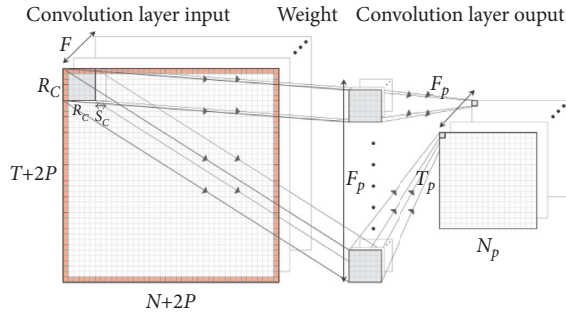


FIGURE 3: Convolution operation diagram [10].

adjacent level are compulsively connected with each other and have correlation relations. The input filtering function of this structure in local space has strong adaptability. It mainly includes data input layer and hidden layer, and each hidden layer is subdivided into convolution layer, sampling layer, and pooling layer. The core of the algorithm is convolution and pooling. The supervised learning mode is used to construct the training network, which mainly includes the preceding training propagation and reverse training propagation. The forward training propagation changes the input training sample matrix in layers, and the output of each layer is the input of the next layer, and finally the convolutional feature matrix is extracted [13, 14].

Set the width of the input data as N and the height as T ; then the width and height of the output image can be calculated by the following formula:

$$N_p = \frac{N + 2P - R_c}{S_c} + 1, T_p = \frac{T + 2P - R_c}{S_c} + 1. \quad (2)$$

Then the operation can be expressed as

$$a_{f,l,m}^{(i)} = \sum_{f'=0}^{F_c-1} \sum_{j=0}^{R_c-1} \sum_{k=0}^{R_c-1} w_{f',j,k}^{(v)f} h_{f',s_c l+j,s_c m+k}^{(i)(v)}, \quad (3)$$

where $v \in [0, N-1]$, $f \in [0, F_{v+1}-1]$, $l \in [0, N_{v+1}-1]$, and $m \in [0, T_{v+1}-1]$. After each hidden layer, an activation function g is required to introduce nonlinear features. After counting completion operation, the input of the next layer is $h_{f,l+p,m+p}^{(i)(v+1)} = g(a_{f,l,m}^{(i)(v)})$.

2.4. Pooling. Figure 4 is the pooling operation diagram. Figure 4 is reproduced from the work of Chen and Tang (2021) [10] [under the Creative Commons [Attribution License/public domain]]. In convolutional neural networks, pooling layer is often used to further adjust the output of the convolutional layer. After extracting object features through the convolutional layer, pooling layer operation usually needs to be inserted between two adjacent convolutional layers. Pooling layer can further reduce the number of neuron connections between convolutional layers, accelerate the learning speed of the model, and alleviate the problem of overfitting. Specifically, the pooling operation computes the local statistical characteristics of a location to replace the elements of the input data on that location region. Through the operation of the pooling layer, the number of parameters in the entire network is reduced, which alleviates the complexity of parameter calculation and storage [15, 16].

The average pooling operation can be expressed as

$$h_{f,l+p,m+p}^{(i)(v+1)} = a_{f,l,m}^{(i)(v)} = h_{f,s_p l+j+p,s_p m+k+p}^{(i)(v)} k_{(f,l,m)}^{(i)(p)} + p. \quad (4)$$

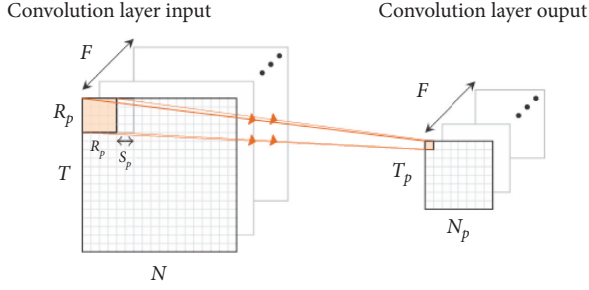


FIGURE 4: Pooling operation diagram [10].

Setting $j_{(f,l,m)}^{(i,p)}$, $k_{(f,l,m)}^{(i,p)}$ represents the pixel index with the largest value in the pooled perception field, and the input of the next layer is got:

$$a_{f,l,m}^{(i,v)} = \sum_{j,k=0}^{R_p-1} h_{f,s_p l+j+p, s_p m+k+p}^{(i,v)} \quad (5)$$

2.5. Convolution to the Full Connection Layer. The learning of several convolutional layers and pooling layers can abstract the information in the input data into features with higher information density; that is to say, the convolutional layer and pooling layer can be regarded as the process of data feature extraction and learning [17]. After completing this step, the full connection layer can be used to complete the classification task. This process can be expressed as

$$a_f^{(i,v)} = \sum_{f'=0}^{F_v-1} \sum_{l=0}^{R_v-1} \sum_{m=0}^{R_v-1} w_{f',l,m}^{(v)f} h_{f',l+p,m+p}^{(i,v)} \quad (6)$$

The next level of input is obtained by activating the function $h_f^{(i,v+1)} = g(a_f^{(i,v)})$. Figure 5 is the one-dimensional output obtained by convolution. Figure 5 is reproduced from the work of Chen and Tang [10] [under the Creative Commons [Attribution License/public domain]].

2.6. Fully Connected Layer. The maximum pooling operation uses the maximum value of the local region as the output result, which can significantly reduce the offset caused by the parameter error of the convolution layer and retain more local feature information of the input data. Common pooling operations include the average value of adjacent rectangular regions and the weighted average function of the distance from the center pixel.

$$a_f^{(i,v)} = \sum_{f'=0}^{F_v-1} w_{f'}^{(v)f} h_{f'}^{(i,v)} \quad (7)$$

2.7. Output Layer. The last layer of the model is the output layer, outputting

$$a_f^{(i,v)} = \sum_{f'=0}^{F_v-1} w_{f'}^{(v)f} h_{f'}^{(i,v)}, \quad (8)$$

$$\text{softmax}(a_f^{(i,(V-1))}) = \frac{e^{a_f^{(i,(V-1))}}}{\sum_{f'=0}^{F_{v-1}-1} e^{a_{f'}^{(i,(V-1))}}} \quad (9)$$

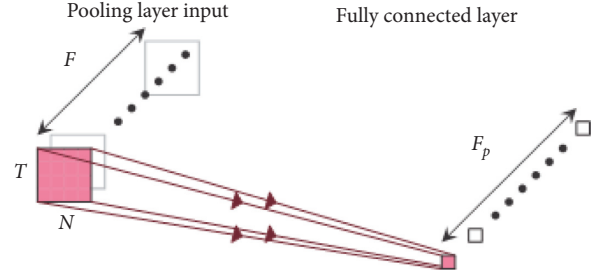


FIGURE 5: One-dimensional output obtained by convolution [10].

where V is the total number of layers of the model. o is the output function, and the following Softmax [18] function is usually used when the model is used to classify tasks.

After obtaining the features of the training sample through convolution, if the features obtained through convolution are directly used, it is inevitable that there will be a large amount of computation. It is very inconvenient to learn a classifier with extremely high-dimensional feature input, and overfitting phenomenon is very easy to occur in this case. Faced with such a problem, the solution is the feature after convolution, which may be applicable to different regions of the sample. In order to describe the sample with a large dimension, aggregative statistics are a solution for the feature of different locations. People can calculate the average (or maximum) of a particular feature of the sample, and the average (or maximum) is calculated over an area of the sample. This will reduce the dimension, and the overfitting phenomenon will not easily occur. The operation of this aggregation is called pooling. In a different sense, pooling can also be called downsampling.

DL is one of the powerful tools in the current wave of artificial intelligence. Its essence is a rule learning method based on multilayer neural network, which takes massive data as input. The neurons in the input layer receive the input data and transmit the data to the neurons in the hidden layer. The neurons in the hidden layer process the data with activation function and transmit the processed data to the output layer. Finally, neurons in the output layer output the results. In DNN, the more hidden layers, the more complex the structure. At this time, the parameters of the whole network become more, and the abstract representation of each layer is deeper than that of the previous layer, so DNN has more powerful performance. With the booming development of DL, DNN has also been introduced into the field of communication to solve traditional communication problems such as modulation identification. The modulation recognition technology based on DL can give full play to the performance advantages of DNN and make up for the shortcomings of the original algorithms [19].

The deep confidence network includes hidden layer and visible layer, and the fully connected mode is adopted between layers. The model structure of CDBM is shown in Figure 6.

The restricted Boltzmann mechanism is used in DBN sample model training to form neural network perception, and its model is shown as

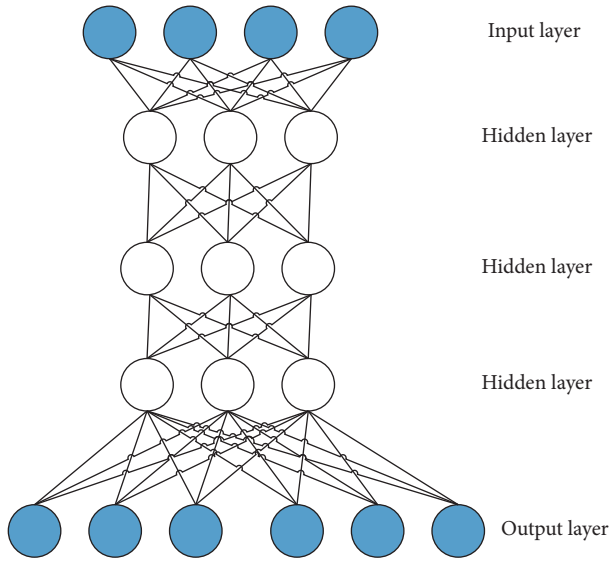


FIGURE 6: The model structure of CDBM.

$$E(v, h, \theta) = - \sum_{i=1}^m \sum_{j=1}^n v_i h_j \omega_{ij} - \sum_{i=1}^m d_i v_i - \sum_{i=1}^m c_i h_i, \quad (10)$$

where v_i and c_i represent the bias of the i -th display layer and its corresponding display neuron, respectively; ω_{ij} represents the weight value of neuron connection between the display layer and the hidden layer; h_i represents the j -th hidden layer; corresponding d_i represents the paranoid value of neurons in the hidden layer; $\theta = \{\omega_{ij}, c_i, d_i\}$.

The calculation method of weight value is

$$\omega_{ij} = \eta [E_{\text{data}}(v_i h_j) - E_{\text{mod}}(v_i h_j)], \quad (11)$$

where i and j represent the number of visible layer nodes and the number of hidden layer nodes, respectively; η represents the efficiency value of learning; simultaneous v_i and h_j represent binary variables; E_{data} and E_{mod} represent the expected value of the training sample and the output sample, respectively.

3. Feature Learning Algorithm Based on CNN&DBN to Extract Music Score

With the development of Internet and digital audio technology, music information retrieval has gradually become a research hotspot. The effective recognition and classification of music genres are an important research direction. It is also a hot research direction to identify the type of musical instrument and determine the specific name of musical instrument in the unknown music segment [20]. The correct extraction of musical features is an important indicator for the classification of musical factions. The vast majority of music songs contain the corresponding musical instrument audio, so the effective recognition and classification of musical instruments can provide strong support for music information retrieval. Therefore, the recognition and classification of musical instruments are also an important part of music information retrieval, which is of great significance

to music retrieval. For people, the recognition and classification of musical instruments are relatively easy to complete, as long as the person has a strong musical accomplishment; it can be more accurate recognition and classification of musical instruments, but most of the ordinary music audience does not have this ability, so it is necessary to teach the computer how to automatically identify and classify musical instruments. From an acoustic point of view, the timbre of an instrument is the main basis for distinguishing one instrument from another, accurately marking the unique characteristics of each instrument [21–23]. The timbre of a musical instrument is mainly determined by the vibration of the articulating part of the instrument. Different vibration states lead to different overtones and waveforms, and the proportion of harmonics in overtones determines the timbre of musical instruments. However, the proportion of overtones varies widely, and the different playing techniques used for the same instrument can also show significant changes in timbre, which can also be mistaken for the sound produced by other instruments. This will also lead to difficulties in the feature extraction of musical instruments from musical signals, which will lead to low accuracy in the recognition and classification of musical instruments. There are few research works on musical instrument recognition and classification.

According to the background of the continuum hypothesis and its analogy to the sound signal, it can be seen that the musical information is composed of a large number of sampling points, and there are gaps in time between the sampling points. Therefore, the musical information composed of discrete sampling points can be regarded as composed of numerous frames distributed continuously without gaps. Since the speech signal is a nonstationary signal, it is meaningless to directly extract all the features of a complete audio file, but the audio changes slowly in a short time and can be considered short-term stationary. At this point, the assumption of stability continuity allows us to estimate statistical information from observational data. This short stationary partition is the signal frame. On the whole, the characteristic parameters of music signal change with time, so it belongs to a nonstationary state and cannot be analyzed by the relevant techniques for processing stationary signals. However, considering that the voice in the song is formed by the movement of people's oral muscles, this kind of oral muscle movement is relatively slow and has a lag. At the same time, the accompaniment in the song is formed by the process of percussion and string vibration of piano, guitar, and other instruments, so it also has a certain lag. From this point of view, although the music signal has the time-varying characteristics mentioned above, it can still be regarded as a quasi-static process by extracting a certain short-time interval of the signal; that is, it shows that the music signal has short-time stability.

The advantages of DBN model are that it overcomes the problems of neural network algorithm, such as high requirement on data, very slow convergence rate, and poor global optimization of local solution. Combined with the advantages of learning algorithm, this paper proposes CNN & DBN algorithm, which uses the characteristics of learning

algorithm to preprocess and extract features of each musical instrument music sample score [24]. Then it is input into the deep confidence network, followed by adjusting the parameter to predict the types of test instruments, and, through repeated comparison test and tuning, the highest recognition accuracy is finally reached. The structure diagram of feature learning algorithm of CNN&DBN is shown in Figure 7.

The detailed steps are as follows:

- (1) Firstly, the input dataset is preprocessed, mainly through the pitch feature matrix and constant changes to complete noise filtering and volume calibration. After preprocessing, the sample dataset of music soundtrack is obtained.
- (2) The improved music score training model based on CNN model is adopted to conduct supervisory learning training on samples and generate training feature sample model.
- (3) The introduction of the original test sample set also requires noise reduction for the test set. After processing, the feature sample model extracted by CNN was input into the test combination. New feature vectors are extracted.
- (4) The extracted feature vector is combined with the classification set Y of the original sample set to generate a new input sample set, and the sample set is input into the DBN model as the input set for training, and the final classification training result is obtained.

Based on the research of the new model and the polyphony combination of multiple parts of the score, the algorithm tells CNN to adjust the model algorithm adaptively and introduces the simulated human auditory attention to establish the music score recognition classification benchmark model. The key of the model is the automatic filtering due to the difference of brain structure. When people hear music, they will focus on the rhythm and main sound parts of the music so as to identify different instruments. Therefore, this model modeling method proposed in this paper simulates the operation process of human brain. Combined with the characteristics of background musical instruments, attention model is introduced to set corresponding feature weights for different musical instrument bands corresponding to key components [25]. The flow chart of BNN model for music recognition and classification based on focus network is shown in Figure 8.

As shown in Figure 8, BNN model based on concern network is divided into convolution layer, batch standardization layer, rule function layer, and maximum pooling layer. Among them, the pooling layer is part of attention network and convolution layer, so as to generate feature vectors with weight values, and Sigmoid function is used to obtain recognition feature vectors. The attention network includes the convolution, the whole connection layer and the corresponding normalization function. After the convolution of the characteristic matrix r_n , it becomes a medium-sized matrix. Each medium-sized matrix input connection

layer will get an appropriate amount of attention $v_n = f_{\text{all}}(r_n; \theta_{\text{all}})$, set θ_{all} as the parameter of all connection layers, and each component of all attention vectors uses the normalization function to normalize the memory, and the sum a_n exceeds the normalization amount M . The attention weight value ∂n is obtained, as shown in the two following equations:

$$a_n = \sum_{m=1}^M \frac{\exp(v_{n,m})}{\sum_{n=1}^N \exp(v_{n,m})}, \quad (12)$$

$$\partial n = \frac{\exp(a_n)}{\sum_{n=1}^N \exp(a_n)}, \sum_{n=1}^N \partial n = 1. \quad (13)$$

The flow of BNN model based on concerned network is similar to that of traditional neural network algorithm [25]. The core lies in the attraction of attention factor of convolution layer. The number of operations and space occupation of single-layer convolution in this algorithm are shown in the two following equations:

$$T \sim O\left(\sum_{i=1}^D M_i^2 \cdot K_i^2 \cdot C_{l-1} \cdot C_l\right), \quad (14)$$

$$S \sim O\left(\sum_{i=1}^D K_i^2 \cdot C_{l-1} \cdot C_l + \sum_{i=1}^D M_i^2 \cdot C_l\right), \quad (15)$$

where M represents the length of the convolution kernel corresponding to the eigenvalue output by the convolution kernel in the convolution layer; C_{l-1} represents the input channel value corresponding to the upper layer of the convolution kernel; C_l represents the output channel value of this layer; M_i^2 represents the area of the eigenvector; K_i^2 represents the calculated area of the convolution kernel; D represents the depth of the network. Thus, the value of the channel is the depth of the convolution. Therefore, the time complexity of the convolution algorithm is obviously positively correlated with the area of the convolution layer and the feature vector [26–28].

4. Example Verification and Analysis

The data samples selected in this paper are 3 seconds long and contain music clips of various musical Instruments, while the musical instrument training samples contain 3120 samples of musical instrument, as well as a total of 21,840 music samples. Musical instruments selected 670 test samples, with a total of 4690 test samples. In the pretreatment process, it is preferred to add labels after noise reduction to the samples. The model of the training set is used to cross process the test set and the training set. At the same time, it also verifies the prediction ability of the score of the test samples. Finally, the predicted score labels are compared with the actual score classification. The final accuracy is the result of the average calculated value. The attribute distribution diagram of feature samples is shown in Figure 9.

Combined with the research on the performance of the traditional classification algorithms, it was found that deep

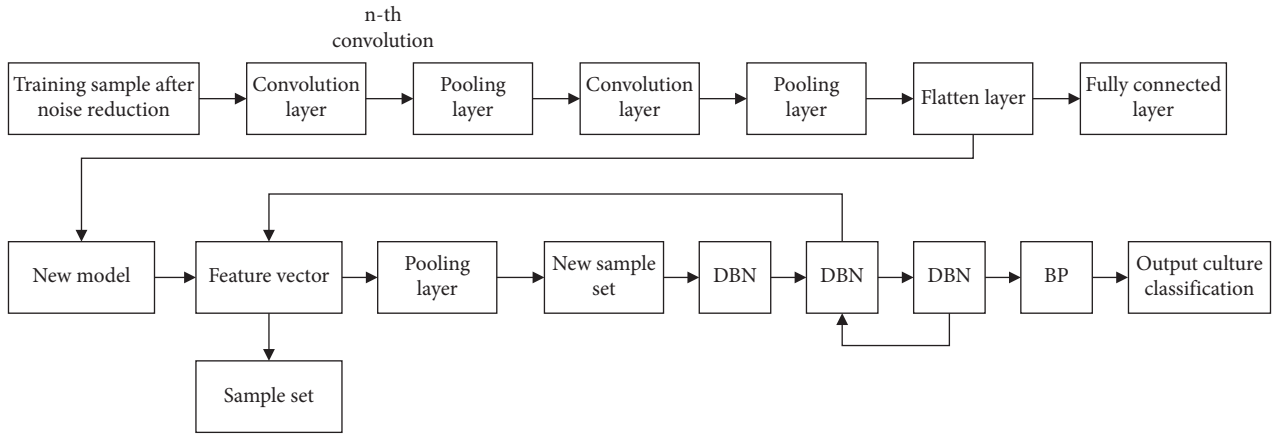


FIGURE 7: Structure diagram of feature learning algorithm of CNN&DBN.

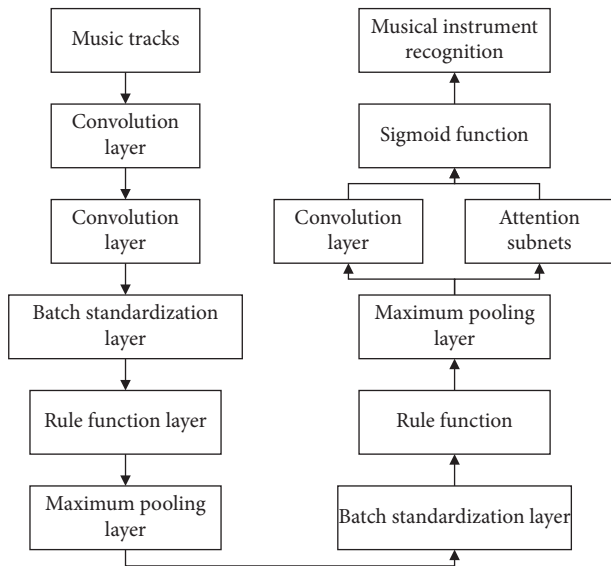


FIGURE 8: Flow chart of BNN model for music recognition and classification based on focus network.

learning sample training process requires tremendous force and execution time is longer, and the conventional experimental environment is extremely easy to interrupt the experiment process and system outage; the unit usually needs weeks or even months of comparative experiments to achieve the basic state of stable operation. Other algorithms will also be compared [29]. This paper adopts the finite cycle method to carry out the experiment, limiting the execution cycle and the number of cycles. After repeated calculation, 9 cycles are selected to be executed. The accuracy comparison experiment between the CNN&DBN model tested in this paper and the decision tree is carried out. As well as the change of the execution accuracy, the experimental results show that the performance is best when the number of cycles is 9. The contrast of confusion matrix is shown in Figure 10. Compared with the average accuracies of other classical algorithms, the algorithm in this paper combines the advantages of CNN in feature

extraction and the high execution efficiency of DBN algorithm and introduces attention weight into CNN algorithm. Therefore, both accuracy and performance are greatly improved, which are significantly higher than those of other algorithms.

The music recognition algorithm based on CNN&DBN proposed in this paper has significantly improved the recognition results applied to the same dataset in terms of various scores and overall recognition accuracy compared with other classical models, especially the recognition improvement effect of violin is more obvious. Compared with other models, the overall music recognition of this model is more uniform, which effectively improves the problem of uneven recognition effect between different categories. By changing the proportion of labeled samples in the training set of traditional Chinese musical instruments for network fine-tuning training [30], the training set is constant, and the recognition accuracy of musical instruments increases with the increase of the number of labeled samples in network fine-tuning. However, the accuracy of the algorithm can still reach more than 90%. In this way, when a large number of Chinese traditional musical instrument music samples are identified and classified by deep confidence network, only a small number of samples need to be labeled, which greatly reduces the workload and is conducive to improving the efficiency of identification and classification. The parameter optimization graph of deep learning network is shown in Figure 11.

By comparing the results in the figure, we find that the first hidden layer has the greatest influence on the predicted results, and when the number of nodes in the first hidden layer is about 3 times of the characteristic dimension of the input sample, the performance of the network converges basically [31]. Meanwhile, as the number of neuron nodes in each layer increases, the training duration will also increase. Therefore, except for the first layer, the number of neuron nodes in other hidden layers will be less appropriately selected. Taking the 2s-long Chinese traditional musical instrument music fragment as a sample, the training set contains 2340 samples for each type of musical instrument, with a total of 14,040 musical instrument samples. The

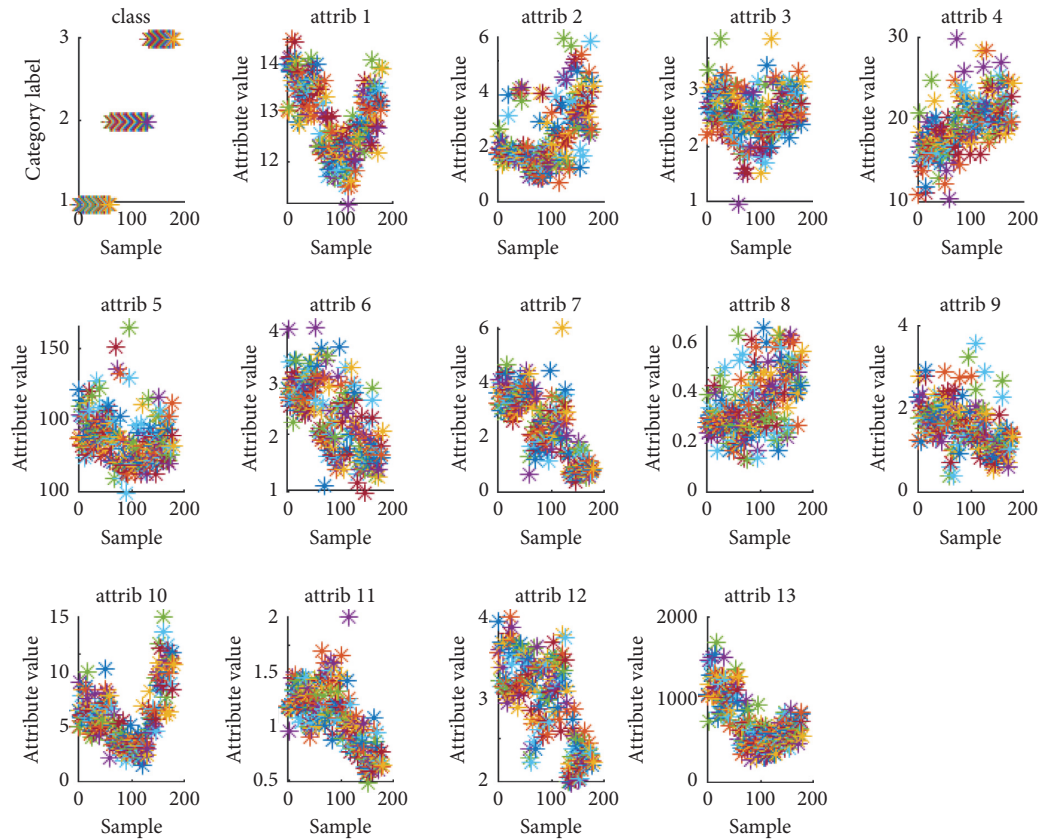


FIGURE 9: Attribute distribution diagram of feature samples.

validation set and test set contain 780 musical instrument samples per category, for a total of 4,680 musical instrument samples. First, all musical instrument samples were labeled, and the training set was used to train the model. Then, cross-validation was carried out on the validation set, and the musical instrument samples in the test set were predicted. Finally, the average recognition and classification accuracy was obtained by comparing the predicted musical instrument labels with the actual musical instrument labels. The characteristic curves for different instruments are shown in Figure 12.

The recognition and classification effect of MFCC features directly input into the traditional classical

classifier is poor, and the lowest recognition and classification accuracy of decision tree is 83.9%, and the k -nearest neighbor classification algorithm (KNN) has the highest average classification accuracy of 92.7%. The accuracy rate is even as high as 99.2%, which is 6.5% higher than that of KNN. The experimental results prove that the music samples of traditional Chinese musical instruments can have better recognition and classification effect after further learning by deep confidence network. By comparing the recognition rate confusion matrices of the six classification methods, it is found that different classification methods have different recognition effects on various traditional musical instruments, which may be

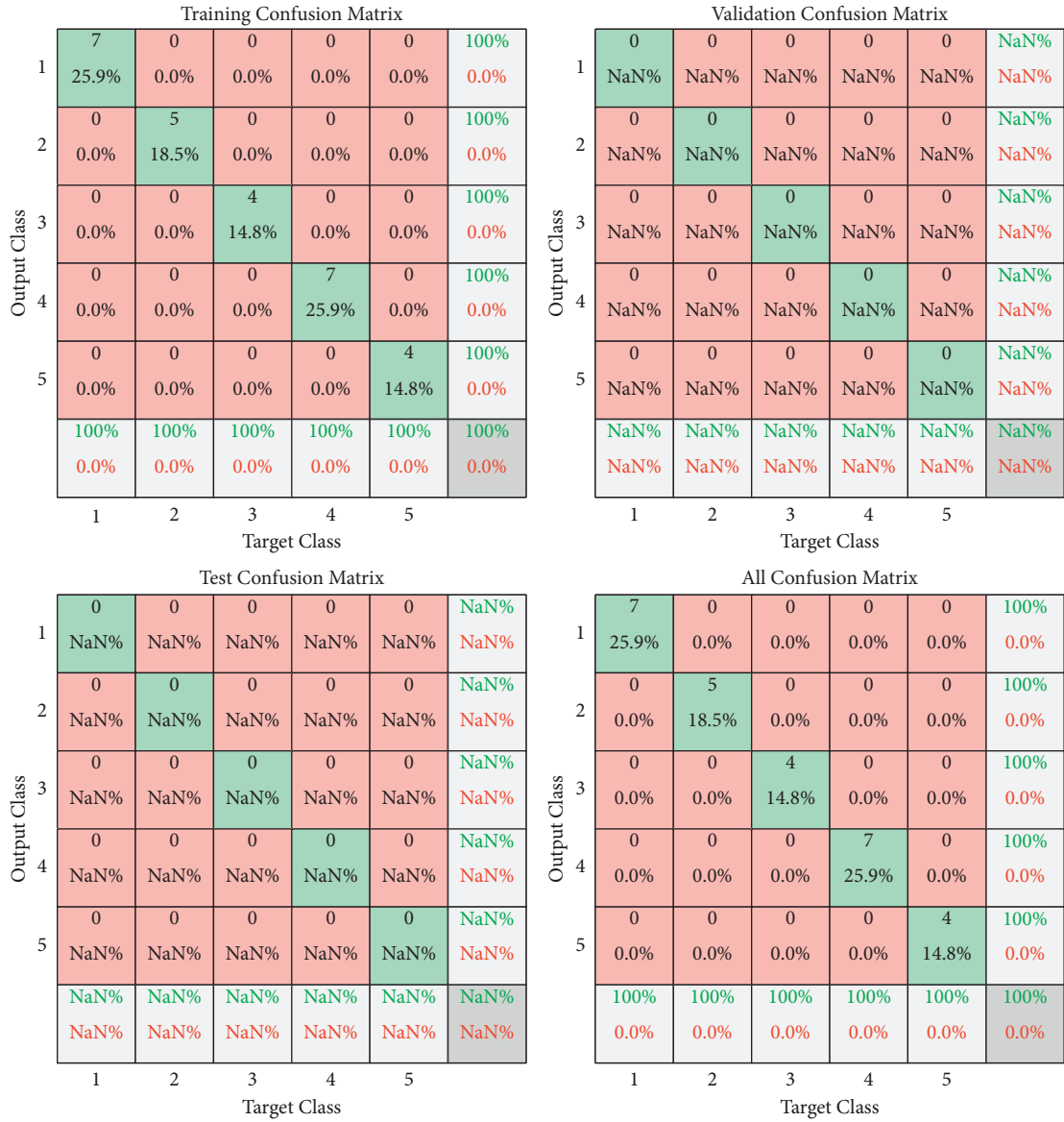


FIGURE 10: Contrast of confusion matrix.

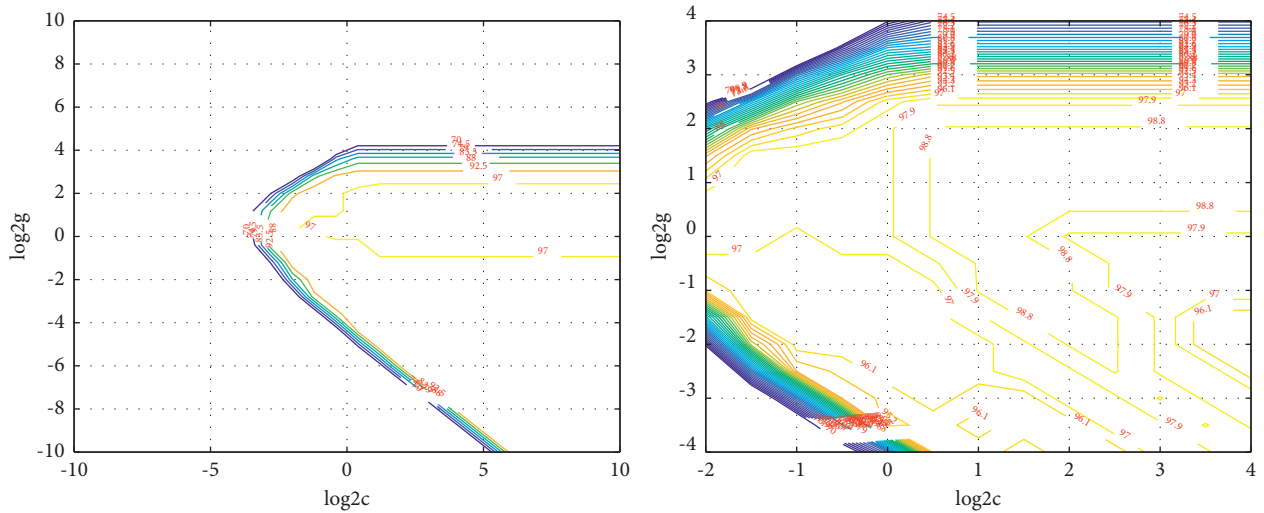


FIGURE 11: Parameter optimization graph of deep learning network.

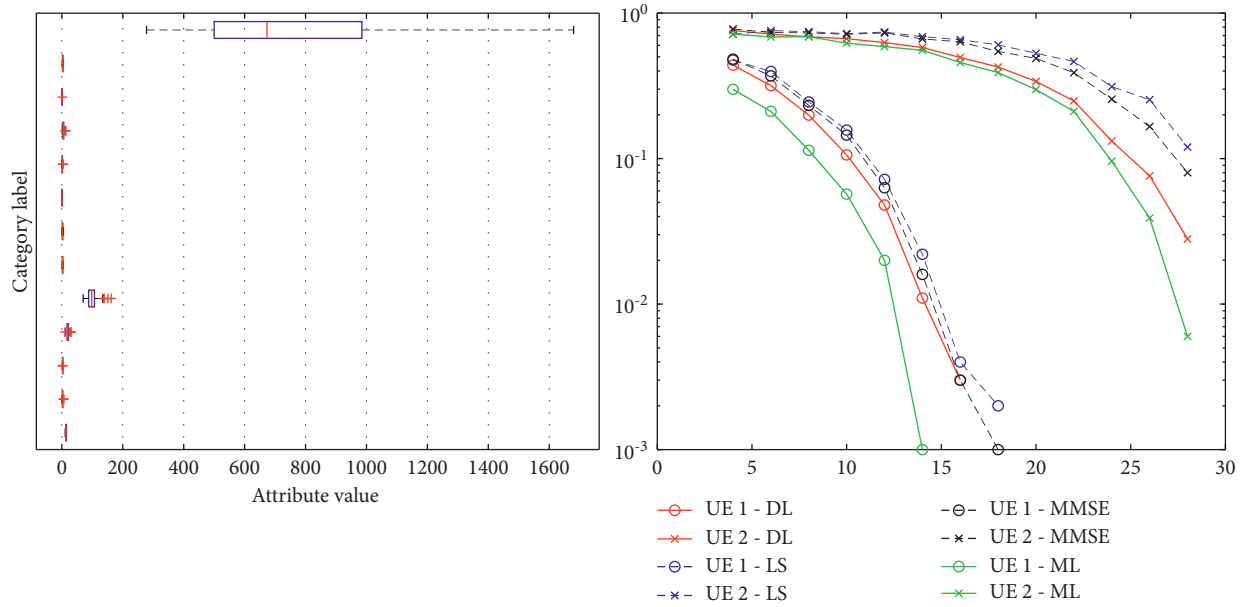


FIGURE 12: Characteristic curves for different instruments.

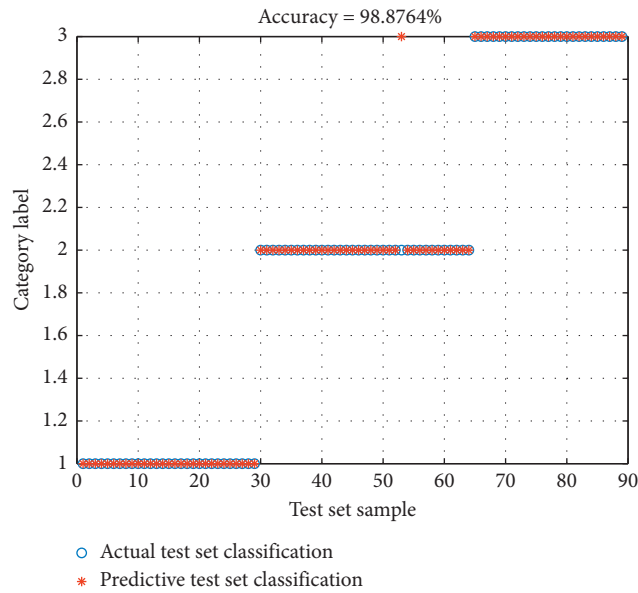


FIGURE 13: Classification accuracy of the classifier.

TABLE 1: Comparison between algorithms for each score recognition result.

Method	Piano	Violin	Viola	Bass	Saxophone	Xylophone
Decision tree	0.899	0.901	0.892	0.879	0.832	0.836
KNN	0.891	0.894	0.835	0.887	0.875	0.841
SVM	0.902	0.876	0.894	0.908	0.894	0.855
CNN&DBN	0.913	0.945	0.931	0.903	0.903	0.863

related to the training of the inherent characteristics of musical instruments in these classifiers. In traditional identification method, erhu and flute recognition rate is low, but, with deep belief networks of instrument samples

after further training, the two kinds of musical instrument recognition rate also have the obvious improvement, which proves the recognition and classification task of deep belief network for traditional instruments. The

advantage of using this method is that more musical features of the essence of traditional instruments can be extracted. The classification accuracy of the classifier is shown in Figure 13. Table 1 shows comparison between algorithms for each score recognition result.

5. Conclusion

With the continuous development of machine learning, more and more practical application problems have been put forward, among which the recognition of music score is still a challenge. In this paper, convolutional neural network (CNN) is improved to identify and extract feature vectors of music soundtrack, and the high efficiency of the implementation of deep confidence network (DBN) is introduced. Taking the feature vector set extracted by CNN as DBN input set, a feature learning algorithm based on CNN-DBN was established to extract music score. Through the comparison of experiments, it is found that the model proposed in this paper shows more accurate recognition ability and good execution performance in the recognition ability of multiple types of polyphony music. The recognition rate of the improved algorithm applied to music recognition is as high as 98.4%, which is obviously better than those of other classical algorithms. The experimental data fully demonstrate the obvious advantages of the proposed algorithm in the recognition of music soundtrack and further prove the superiority of using the deep confidence network to identify and classify musical instruments.

Although the research in this paper has achieved some results compared with the traditional methods, due to the limitations of my theoretical level and research time, there are still many deficiencies in this paper which need to be improved. The following aspects need to be improved in the follow-up research work.

Firstly, in the Internet information resources, the information base of audio retrieval is very small. How to make the retrieval system adapt to the massive music information retrieval needs further study.

Secondly, the database capacity of the two experiments in this paper is relatively small, which only contains tens of thousands of music samples. In the face of massive music data on the network, the computational complexity of training is very high, and the hardware may not be able to afford it. We should try to learn how to do parallel processing and how to implement it with GPU, which is also a direction of future research.

Third, in the experiments based on HMM and CDBN model, the retrieval recognition rate for melody is not good enough, which is difficult to make many people able to better obtain the music information they need, so in the future music retrieval system in this aspect of the related technology needs to make breakthroughs.

Fourth, convolutional deep confidence network is time-consuming in training, and there is no unified theoretical guidance for network structural parameters. Therefore, how to set network structural parameters adaptively according to different samples is worth discussing.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] Q. Vo, G. Lee, S. Kim, and H. Yang, "Recognition of music scores with non-linear distortions in mobile devices," *Multimedia Tools and Applications*, vol. 77, no. 12, pp. 15951–15969, 2018.
- [2] L. Mengarelli, B. Kostiuk, J. G. Vitório, and M. A. W. C. N. Tibola, "OMR metrics and evaluation: a systematic review," *Multimedia Tools and Applications*, vol. 79, no. 9–10, pp. 6383–6408, 2020.
- [3] M. Ambrosiano, M. T. Bevacqua, T. Isernia, and V. Pascazio, "The tomographic approach to ground-penetrating radar for underground exploration and monitoring: a more user-friendly and unconventional method for subsurface investigation," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 62–73, 2019.
- [4] D. C. Bowden, K. Sager, A. Fichtner, and M. Chmiel, "Connecting beamforming and kernel-based noise source inversion," *Geophysical Journal International*, vol. 224, no. 3, pp. 1607–1620, 2020.
- [5] A. T. Bukkapatnam, P. Depalle, and M. M. Wanderley, "Defining a vibrotactile toolkit for digital musical instruments: characterizing voice coil actuators, effects of loading, and equalization of the frequency response," *Journal on Multimodal User Interfaces*, vol. 14, no. 3, pp. 285–301, 2020.
- [6] R. H. Mu and X. Q. Zeng, "A review of deep learning research [J]," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 4, pp. 1738–1764, 2019.
- [7] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016.
- [8] H. Chen, Y. Zhang, Y. Cao, and J. Xie, "Security issues and defensive approaches in deep learning frameworks," *Tsinghua Science and Technology*, vol. 26, no. 6, pp. 894–905, 2021.
- [9] J. Hu, J. Chen, L. Zhang, and Y. Q. H. C. Liu, "A memory-related vulnerability detection approach based on vulnerability features," *Tsinghua Science and Technology*, vol. 25, no. 5, pp. 604–613, 2020.
- [10] Yu Chen and Z. Tang, "Research on the construction of intelligent community emergency service platform based on convolutional neural network," *Scientific Programming*, vol. 2021, p. 5089236, 2021.
- [11] X. Yan, S. Hu, Y. Mao, and Y. H. Ye, "Deep multi-view learning methods: a review," *Neurocomputing*, vol. 448, pp. 106–129, 2021.
- [12] Z. Sadeghi, "Deep learning and developmental learning: emergence of fine-to-coarse conceptual categories at layers of deep belief network," *Perception*, vol. 45, no. 9, pp. 1036–1045, 2016.
- [13] S. Correa and F. R. Mc Causland, "Leveraging deep learning to improve safety of outpatient hemodialysis," *Clinical Journal of the American Society of Nephrology*, vol. 16, no. 3, pp. 343–344, 2021.

- [14] K. D. Song, "Current status of deep learning applications in abdominal ultrasonography," *Ultrasonography*, vol. 40, no. 2, pp. 177–182, 2021.
- [15] A. Ghods and D. J. Cook, "A survey of deep network techniques all classifiers can adopt," *Data Mining and Knowledge Discovery*, vol. 35, no. 1, pp. 46–87, 2021.
- [16] S. Do, K. D. Song, and J. W. Chung, "Basics of deep learning: a radiologist's guide to understanding published radiology articles on deep learning," *Korean Journal of Radiology*, vol. 21, no. 1, p. 33, 2020.
- [17] M. Du, F. Yang, N. Zou, and X. Hu, "Fairness in deep learning: a computational perspective," *IEEE Intelligent Systems*, vol. 36, no. 4, pp. 25–34, 2021.
- [18] M. A. Talab, H. Tao, and A. A. M. Al-Saffar, "Review on deep learning-based face analysis," *Advanced Science Letters*, vol. 24, no. 10, pp. 7630–7635, 2018.
- [19] C. M. Stultz, "The advent of clinically useful deep learning[J]," *JACC-Clinical Electrophysiology*, vol. 5, no. 5, pp. 576–578, 2019.
- [20] L. Pehrson, C. Lauridsen, and M. Nielsen, *Ultraschall in der Medizin - European Journal of Ultrasound*, vol. 39, no. 04, pp. 379–381, 2018.
- [21] A. Jonsson, "Deep reinforcement learning in medicine," *Kidney Disease*, vol. 5, no. 1, pp. 18–22, 2019.
- [22] Y. Chen, Y. Wang, Y. Gu, and X. P. X. He, "Deep learning ensemble for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 6, pp. 1882–1897, 2019.
- [23] H. Li, "Deep learning for natural language processing: advantages and challenges," *National Science Review*, vol. 5, no. 1, pp. 24–26, 2018.
- [24] H. Tembine, "Deep learning meets game theory: bregman-based algorithms for interactive deep generative adversarial networks," *IEEE Transactions on Cybernetics*, vol. 50, no. 3, pp. 1132–1145, 2020.
- [25] Z. Zhang, Y. Zhao, X. Liao, and W. K. Q. S. Shi, "Deep learning in omics: a survey and guideline," *Briefings in Functional Genomics*, vol. 18, no. 1, pp. 41–57, 2019.
- [26] D. Lee, J. Lee, J. Ko, and J. K. Y. Yoon, "Deep learning in MR image processing," *Investigative Magnetic Resonance Imaging*, vol. 23, no. 2, p. 81, 2019.
- [27] C. He, A. Cao, J. Chen, and X. W. J. R. Liang, "Direction finding by time-modulated linear array," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 7, pp. 3642–3652, 2018.
- [28] Z. Jamil, A. Jamil, and M. Majid, "Artifact removal from EEG signals recorded in non-restricted environment," *Bio-cybernetics and Biomedical Engineering*, vol. 41, no. 2, pp. 503–515, 2021.
- [29] M. A. Jatoi and N. Kamel, "Brain source localization using reduced EEG sensors," *Signal, Image and Video Processing*, vol. 12, no. 8, pp. 1447–1454, 2018.
- [30] G. J. Mendis, J. Wei-Kocsis, and A. Madanayake, "Deep learning based radio-signal identification with hardware design," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 5, pp. 2516–2531, 2019.
- [31] S. L. Fowler, H. Calhoun, and A. D. Warner-Czyz, "Music perception and speech-in-noise skills of typical hearing and cochlear implant listeners," *American Journal of Audiology*, vol. 30, no. 1, pp. 170–181, 2021.

Research Article

Smart Home Product Layout Design Method Based on Real-Number Coding Genetic Algorithm

Nan Jiang 

Department of Space Design, Kookmin University, Seoul 02707, Republic of Korea

Correspondence should be addressed to Nan Jiang; jiangnan@kookmin.ac.kr

Received 10 June 2022; Accepted 22 June 2022; Published 6 July 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Nan Jiang. 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 problems of poor layout design efficiency of smart home products and low rationality of layout planning, a layout design method for smart home products based on a real-number coding genetic algorithm is proposed. The principles of smart home product layout design are analyzed and the smart home system architecture based on the Internet of Things is designed. Divide the combination of smart home product layout space according to spatial function, extract the visual features of smart home product layout, build a smart home product layout optimization model based on the two constraints of the total area and individual area of smart home product spatial layout, and design a real-coded genetic algorithm. The model is solved to improve the global convergence of the algorithm, and the optimization method for the layout of smart home products is obtained. The experimental results show that the layout design method of smart home products based on a real-number coding genetic algorithm can accurately extract the layout features of smart home products and accurately classify the number of pairs of homes. The efficiency of layout design and rationality of layout planning of smart home products are better, which reflects the effectiveness of this method.

1. Introduction

In the era of rapid development of information and economy, people's requirements for the home environment are constantly improving. Especially in the design of home furnishing and furnishings, the transformation from function-based to more beautiful and practical value, if the hard decoration is the "skeleton" of the building, the soft furnishing and furnishings of the home are the "flesh and blood" essence of its space part. The embodiment of modern home furnishings in the interior space presents a variety of designs with different styles [1]. The smart home is based on housing as a platform, using integrated wiring technology, communication technology, automatic control technology, and other related technologies to analyze the communication mode between the home and outside. The extensive use of smart home can effectively improve the quality of life. Multifunctional furniture can achieve different functions in a narrow range and achieve the purpose of flexible

transformation and space-saving. Ensure the accuracy, immediacy, and stability of communication between wireless home systems, better meet the multilevel management needs of smart home monitoring, remote control, communication, etc., and make more contributions to the improvement of residential intelligence [2]. With the improvement of living standards, people put forward higher requirements for the quality of home furnishing. For people with no experience in home design, it is usually a matter of trial and error to actually move the furniture around until you are satisfied. However, this restricts the user's free choice and combination of furniture, and the obtained results often fail to meet the functional and aesthetic requirements [3]. Therefore, studying the layout design method of smart home products has become an important research content in the current related research fields.

Reference [4] discusses the active interaction design strategy of conversational agents in smart home scenarios, discusses the characteristics of initiative from the perspective

of decision-making rights and communication methods, and builds four active styles (straight advisor, blunt advice) on this basis: decision maker, euphemistic suggester, euphemistic decision maker. Through the “Wizard of Oz” experimental method to explore the influence of a user’s gender, age, experience, and living status on active style preference, the optimization scheme of active interaction design of conversational agents in smart home scenarios is obtained. However, this method is mainly aimed at the interactive design of conversational agents in smart home products. The rationality of the layout of different types of smart home products needs to be further analyzed. Reference [5] designed a smart home system based on 2.4 G and ZigBee. The system consists of a cloud server, gateway node, 2.4 G remote control node, and ZigBee sensor network. The gateway uploads the data collected by the node to the cloud server and mobile App display through Wi-Fi, realizes the remote control of each ZigBee node through the mobile App or 2.4 G remote control module, and realizes functions such as security, environmental monitoring, remote control switch, and the intelligent lighting adjustment. However, this method is mainly aimed at the communication deployment of smart homes without an in-depth analysis of the communication capabilities under different smart home product layouts. At the same time, it is to test its performance according to different layout schemes. Therefore, it can be used as a reference in this paper when it is applied to the layout of smart home products.

In view of the above problems, this paper proposes a layout design method for smart home products based on a real-number coding genetic algorithm. Analyzed the principles of smart home product layout design and designed the smart home system architecture based on the Internet of Things. The combined distribution of the layout space of smart home products is studied, and the total area and individual area of the smart home product space layout are used as constraints to build a smart home product layout optimization model. On this basis, this paper starts with the real-number coding, using the characteristic that the sum of two pairs of chromosomes before and after the hybridization of coding chromosomes remains unchanged, and the maximum value of the offspring is the value taken by the parent, and the minimum value is the value taken by the parent, through the real-number coding genetic algorithm to solve the model, the smart home product layout optimization method is obtained quickly and accurately.

2. Materials and Methods

2.1. Layout Design of Smart Home Products

2.1.1. Characteristics and Principles of Smart Home Product Layout Design

(1) *Features of Smart Home Product Layout Design.* The layout and design of a smart home are an important part of current creative life. Provide residents with a small-area living environment, maximize the use of resource space, and

gradually accelerate the various needs of urban development. The multifunctional smart home system combines automatic control technology, sensor technology, network communication technology, and embedded technology so that users can access and manage home equipment through wireless network technology. Real-time monitoring of indoor equipment, centralized management of access control settings, and improvement of various alarm functions are effective ways to ensure the safe operation of home equipment [6]. The system of a multifunctional smart home can be subdivided into two levels: hardware and software. The hardware includes servers and home equipment, and the software is composed of various embedded software and mobile phone application software. Smart homes can carry out market innovation and planning and express their own attributes in the natural environment and humanistic characteristics. In fact, the design and architecture of a smart home have the following characteristics:

- (1) The important features of the all-around living environment can make the living environment within a small area have complete living functions. In daily life, all things can be integrated into the same environment, and the space structure can be fully utilized so that a small area of space can play infinite possibilities.
- (2) Culture is also an important part of the living environment. According to the different living environments of people, the customs, concepts, and habits show differences. The structure and design of a smart home should meet the basic requirements of cultural characteristics in the living environment and actively absorb the excellent traditional culture. From the perspective of customs, social values, etc., the living environment reflects the localization characteristics.
- (3) Give full play to the convenient features of the living environment. In the process of smart home product layout planning, a modern living environment is designed according to the current situation. According to specific problems, we propose corresponding plans, pay attention to the interdependence between the living environment and occupants, and gradually meet the needs of occupants for social resources. The layout design of smart home products should meet the actual needs and gradually reduce the waste of resources. For the vast majority of occupants to investigate and analyze, and finally get the most scientific design scheme [7].

(2) *Smart Home Product Layout Design Principles.* The location and rational planning of the residential area should pay attention to the harmonious connection between the residential area and the level of economic income and design an optimal smart home product structure. The following principles should be followed in the design:

- (1) The principle of integrity

The rational planning of the living environment must pay attention to and take into account the overall connection of the living environment. Design and rational planning are closely related projects, and the scheme cannot be designed from an isolated perspective. The needs of the occupants are the theme of this design. The design and structure of a smart home must first pay attention to the overall environment, and then the pattern, modeling style, material selection, etc., must respect the aesthetic foundation, rationally plan the overall design of the living environment, and further complete the visual aesthetics and aesthetics, the laws of harmony and unity that life requires.

(2) The principle of comfort

The structure and design of a smart home are important principles to be followed in the performance principle of the living environment. This principle can be applied in smart home planning. The fresh and convenient living environment can meet the material and spiritual requirements of residents in the new era. To meet the basic living convenience, as well as reasonable activity space, we try our best to meet the living needs of residents and improve their aesthetic level of residents. Let every resident have a healthy, comfortable, beautiful, and convenient living environment.

(3) The principle of function as the main

The economic development needs of different households are different, and the design and structure should be analyzed in combination with the actual situation. Adopt a functional design method to promote the efficient and stable development of the intelligent design. For the smart home design concept, the geographical conditions of the residential area should be fully considered. Choose a design scheme with appreciative and practical value and discuss the architecture and design of smart homes in combination with environmental and economic conditions [8]. Reduce costs as much as possible, formulate a basic design scheme based on creative features, and make full use of the advantages of environmental space to design products and structures with more creative features.

2.1.2. Architecture Design of Smart Home System Based on Internet of Things. The Internet of Things can generally be divided into five layers from bottom to top: perception control layer, access layer, Internet layer, service management layer, and application layer.

The main function of the perception control layer is to “perceive” the environmental parameters and the working parameters of the smart home products and to change the working state of the smart home products as needed. These devices all have ZigBee wireless interfaces [9, 10] to communicate with IoT gateways located at the access layer.

The main device in the access layer is the IoT gateway, which is mainly responsible for connecting many terminals in the perception and control layer to the Internet. On the one hand, it communicates with the terminal of the perception control layer through ZigBee or other interfaces and forwards the data sent by the terminal to the server or forwards the remote control command of the server to the terminal, and on the other hand, it has Ethernet, Wi-Fi, or GPRS, etc. This communication interface can access the local area network of the cell to communicate with the remote server [11].

The main devices in the Internet layer are those general-purpose network devices that are responsible for associating the IoT network with the local area network in the cell and then accessing the Internet or directly accessing the computer network of the network operator. The former can be the switch in the user’s home and the network equipment such as the switch or router in the community, and the latter can be the ADSL modem, Cable Modem, wireless router, fiber router, and other equipment. Of course, both include many central office equipment of the operator. The service management layer mainly includes the application server, Web server, and database server.

The application server is responsible for regular communication with each IoT gateway, obtains the data of the sensing and control layer devices through the gateway, and saves it to the database server in time, while the web server [12, 13] is responsible for publishing the data to the Internet for users to view related information remotely through the browser.

The application layer mainly includes various computing devices such as desktop computers, laptop computers, tablet computers, and smartphones. Its main function is to provide users with a man-machine interface that can interact with the system remotely through a web browser or client software.

In summary, the architecture of the smart home system based on the Internet of Things is shown in Figure 1.

The Internet of Things gateway plays an important role in the whole smart home Internet of Things, and it is a bridge connecting various terminal devices and servers. The gateway communicates with the terminal device through the ZigBee interface or serial port to obtain data and forward the data to the server through the Ethernet or Wi-Fi, or GPRS interface. The design of the gateway adopts the idea of modularization, and it is subdivided into three products: an Ethernet interface, a Wi-Fi interface, and a GPRS interface, according to the common Internet access methods in the family. Users can choose suitable products according to their own conditions.

Because the IoT gateway needs to complete many tasks, in order to better coordinate the operation of each task, an embedded operating system, C/OS-II, is used as its software platform. After the introduction of C/OS-II, application design became very simple. The application is divided into seven user tasks and one system task according to what the gateway should do. The priority is assigned according to the importance of the task and whether it has hard real-time performance. The lower the priority value, the higher the

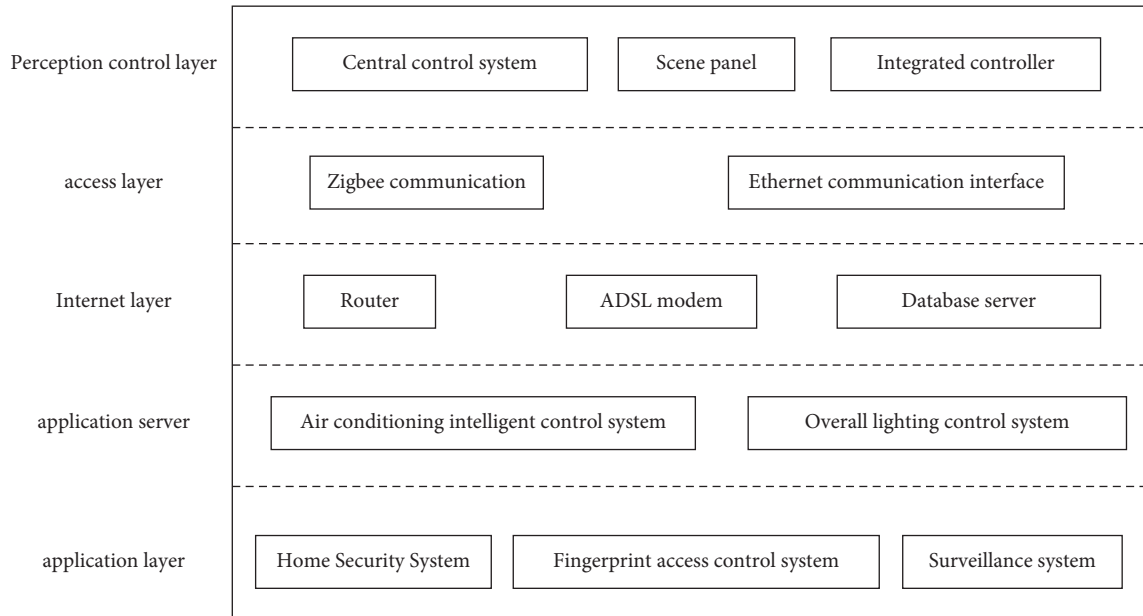


FIGURE 1: The architecture of the smart home system based on the Internet of Things.

priority of the task. The main function of several communication tasks is to process the data received or sent by their corresponding interface chips. The processing of these data is strictly time-limited. If it is not processed in time, it may be corrupted by the data in the next data packet. So, these communication tasks have priorities over other tasks. Among these communication tasks, the Ethernet communication task has the highest priority because the communication rate of Ethernet is the fastest, and the length of the data packet is the longest [14].

One of the main features of the smart home Internet of Things is the ability to connect various smart home products in the house to the Internet, but it is not enough to be able to communicate with devices through the Internet, and the interaction between users and devices must also be realized on the Internet. The application server has stored the device data in the database, and the task of the web server is to display the data on the Internet so that users can view the environmental information and device status in the room through a web browser anytime, anywhere, and can also remotely control the operation of the device state.

2.2. Design Method of Smart Home Product Layout Based on Real-Number Coding Genetic Algorithm. According to the characteristics and principles of smart home product layout design, on the basis of clarifying the smart home system architecture under the application of Internet of Things technology and based on a real-coded genetic algorithm, the smart home product layout design is completed in three parts: the combination of smart home product layout space, the construction of the smart home product layout planning model, and the model solution based on a real-coded genetic algorithm.

2.2.1. Combination of Smart Home Product Layout Space.

According to the function of the space, first locate the overall style, matching colors, and furniture styles; after these are determined, the ceiling, walls, and floors of each interior space interface are decorated and designed. In terms of interior space, it is best to combine soft decoration and hard decoration and use furniture to combine, divide, and contrast the layout and design of the space.

(1) *Combination of Spaces.* Combined interior space is overall planning based on the functional requirements of the building, and the orderly organization from the whole to the single divided space is manifested in the sorting of the size of the space and the household objects that can be accommodated, forming a gap between the indoor space and the home furnishings. The organic connection of the interior space achieves coordination and unity in function and aesthetics, which is also the basis of interior space design. For example, the living room is a space composed of sofa seats, coffee tables, carpets, cabinets, lamps, TV sets, TV cabinets, curtains, and other furniture, providing people with functional areas for leisure, meeting, gathering, and audio-visual. The bedroom can be composed of beds, bedside tables, wardrobes, TVs, dressers, lamps, etc., and mainly provides space for people to rest. There are many forms of combining and expressing various indoor functional spaces, such as axis-symmetrical, centralized combined, radial, and so on.

(2) *Division of Space.* In addition to a certain connection, each space also has its own independence, which is mainly reflected by way of separation. The separation and connection of space is the separation and connection of space in the vertical and horizontal directions. A good separation should be organized in an orderly manner, with appropriate

virtual and real and self-contained systems. This is of great significance to the effect of the entire space design and can present a unique design style. The characteristics of the home furnishing are as follows: Partial separation. Partial separation is the use of furniture such as screens and higher cabinets to block the space, but it does not completely close the space so that the lighting of the space has a breathable feeling which is common in traditional design layouts. Symbolic separation. Symbolic separation is the use of glass, greenery, color, material, height difference, overhangs, and other factors to separate the space, which reflects the effect of continuous separation in space division. This division method is mostly reflected in modern home design. Flexible separation. Elastic separation is the use of folding, lifting, and other activities, such as partitions and curtains, and other furniture, to separate the space, which can be closed or moved according to the requirements of use, and the space can be divided and closed, and the space can be freely stretched. Virtual separation. Virtual separation is the feeling of two spaces formed by different placement or independent design of furniture, carpets, and ceilings in the same space, and the two different functional spaces formed are independent and shared. This design method is usually suitable for small or open layouts.

(3) *Contrast of Space*. The relationship between furniture and interior space should conform to ergonomics, and changes in its size and color also directly affect people's psychological space [15, 16]. The contrast of the space can also be reasonably adjusted by the height, size, and color of the home furnishings so that the space has a well-proportioned sense of rhythm and forms a rhythmic beauty. Designers often use the proportions of furniture and furnishings to integrate them into the design ideas of interior spaces.

2.2.2. Construction of Smart Home Product Layout Planning Model. The smart home product layout approach in this paper allows users to define orientation and distance constraints between furniture. In order to reduce the interaction, the system defaults that all furniture has distance and angle constraints with the room. Therefore, after the user interaction ends, a directed graph of the furniture relationship will be obtained. The nodes of the graph are composed of (1) room and (all) furniture; the edges of the graph indicate that the nodes have a constraint relationship, and the direction indicates the primary and secondary relationship when the user selects, that is, when selecting a furniture pair; the main furniture is selected first, and then the secondary furniture and the directed edge points from the primary furniture to the secondary furniture. Although it is assumed that the user has a certain basic common sense of interior decoration, it does not require the user to make a detailed design of the hierarchical relationship but allows the user to input any furniture pair constraints, so there will be cycles, bidirectional edges, etc., in the furniture relationship directed graph situation.

In order to get a hierarchical tree, it is necessary to delete all the incoming edges of the room node and make it the root

of the hierarchical tree; secondly, check each furniture node so that it has only one incoming edge.

A two-step strategy using a simple size comparison:

Strategy 1: Check all adjacent node pairs; if there are two-way edges, delete a directed edge from the small furniture to the large furniture; if there is only one one-way edge and it is from the small furniture to the large furniture, then flip the direction of the edge.

Strategy 2: Check the neighbors of all incoming edges of each node and keep the incoming edge of the smallest adjacent node.

Establish a resolution multidimensional space block image fusion model for the extraction of visual features of smart home product layout, using block pixel matching method to achieve deep evolutionary learning of visual features of smart home product layout, combined with high-resolution information fusion detection through linear filtering. The pixel reorganization of the edge area realizes the extraction [17, 18] and segmentation of the visual features of the smart home product layout. The segmentation formula is as follows:

$$f = q + \frac{x}{v} \quad (1)$$

Among them, q is the resolution of visual feature extraction of smart home product layout, x is the block time interval parameter, and v is the joint information entropy of feature extraction.

Through the method of two-dimensional parameter fitting, the fitting coefficient of the multidimensional spatial block image for the visual feature extraction of the smart home product layout is obtained as j , and the multilevel visual feature extraction of the smart home product layout is carried out in a single-pixel value distribution area i . By dividing into blocks, the multilevel feature information of the indoor space layout can be obtained as follows:

$$K = i(f \times c) + j(f \times b). \quad (2)$$

According to the multiscale machine learning results, the visual image reorganization of smart home product layout is carried out, and the fuzziness of smart home product layout visual feature extraction is represented by (c, b, m, r) quadruple, y^e, y^r is the entity set for smart home product layout visual feature extraction. Combined with the analysis results of the constraint parameters of the restored image [19], the background value fusion of the visual image of the smart home product layout is realized, and the output is as follows:

$$C = \frac{K(y^e + y^r)}{f} + r + \frac{x}{v}. \quad (3)$$

Select the nearest neighbor function xy group, and the ambiguity distribution set for visual feature detection of smart home product layout is established. From the perspective of feature reuse in the middle layer, the boundary feature components of visual feature detection of smart home product layout are obtained as follows:

$$S = \frac{xy}{C} + r, \quad (4)$$

$$T = \frac{2\pi(g^o + g^u)}{S}.$$

Among them, g^o and g^u are the resolution and information entropy of the visual feature extraction of smart home product layout [20], respectively.

To sum up, build a building interior space layout planning model.

(1) *Objective Function.* Considering the size of the building space, set the objective function of smart home product layout planning as follows:

$$f(x) = \max \sum_{k=1}^K c_k z_k. \quad (5)$$

In the formula, $f(x)$ describes the building interior space layout planning function; c_k represents the number of k -type smart home products in the building interior space; z_k represents the space ratio corresponding to k -type smart home products.

(2) *Constraints.* The total area constraints for the spatial layout of smart home products are described by the following:

$$A = \sum_{k=1}^K Z_K. \quad (6)$$

In the formula, A represents the total area of the interior space of the building.

The individual area constraints of smart home products are described by the following formula:

$$A_j \leq z_j \leq m + \theta. \quad (7)$$

In the formula, A_j represents the minimum demand area of smart home products in the building space; m represents the land area for smart home products; θ represents the planned land area for smart home products.

2.2.3. Model Solving Based on Real-Coded Genetic Algorithm.

The smart home product layout design method based on real-number coding and real-number coding genetic algorithm adopts the real-number coding genetic algorithm [21, 22] to solve the smart home product layout model and realize the optimal design of the smart home product layout. The specific steps are as follows:

- (1) Encoding in the form of 2 rows and R columns.
- (2) Build fitness function $f(m)$:

$$f(m) = \begin{cases} C_{\max} - f(m), & f(m) < C_{\max}, \\ 0, & f(m) \geq C_{\max}. \end{cases} \quad (8)$$

In the formula, C_{\max} is a larger constant.

- (3) Initial population: Most scholars randomly select the initial population, thinking that random selection can traverse all states, but the evolutionary generation of random selection increases, which increases the time required for solving and reduces the efficiency of the method. Based on real-number coding, the smart home product layout model construction method based on a genetic algorithm initializes the population through the following methods [23]:

Step 1: Set a smaller number M , let $i=0$, $f_i = M$;
Step 2: Calculate the fitness value $f_{i(X_{i+1})}$ corresponding to the randomly generated chromosome X_{i+1} , and calculate the average fitness value f_i corresponding to the first $i-1$ chromosomes by the following formula:

$$f_i = \frac{\sum_{i=0}^{i-1} f_i(X_t)}{f_i(X_i)}. \quad (9)$$

Step 3: Determine whether the population accepts individuals through the following process:

- (1) When $f_{i(X_{i+1})} < f_i$, $i+1 < n$, then $X = \{X_1, \dots, X_{i+1}\}$, $i = i+1$, and return to the previous step;
- (2) When $f_{i(X_{i+1})} \geq f_i$, $X = \{X_1, \dots, X_i\}$, go back to the previous step;
- (3) When $f_{i(X_{i+1})} < f_i$, $i+1 = n$, when, $X = \{X_1, \dots, X_{i+1}\}$, go to the next step.

Step 4: End.

- (4) Use the roulette selection method to realize the selection of strategies [24]. The specific steps are as follows:

Step 1: Rank the fitness of individuals in the population in descending order to obtain the sequence $\{X_{N1}, \dots, X_{Ni}, X_{Nn}\}$ of the sorted chromosomes.

Step 2: Use the probability sequence as the basis and basis in the roulette selection process, determine the real number α , obtain n numbers through the following, and use it as the basis in the selection probability process:

$$p_i = \prod_{i=1}^i \alpha_i. \quad (10)$$

Step 3: Normalize the selection probability and calculate the selection probability of each individual by the following formula:

$$p_i = \frac{P_i}{\sum_{i=1}^n P_i}. \quad (11)$$

Step 4: Calculate the cumulative probability q_i corresponding to the chromosome in the next generation population:

$$q_i = \sum_{i=1}^i p_i. \quad (12)$$

Step 5: Select individuals in the population to enter the next generation. When $q_i \leq r \leq q_{i+1}$, chromosome X_{Ni} can enter the next generation population.

- (5) Crossover operation, the crossover operation is realized by the following formula:

$$p_c = \frac{t(p_{\max} - p_{\min})}{T}. \quad (13)$$

In the formula, p_c represents the crossover probability; p_{\max} and p_{\min} represent the maximum crossover probability and the initial crossover probability, respectively.

- (6) Mutation operation, in the process of group evolution, the mutation probability is adjusted and calculated by the following formula:

$$p_m = \frac{p_{m1} + p_{m2}}{2}. \quad (14)$$

- (7) Termination condition, the construction method of smart home product layout model, based on real-number coding genetic algorithm adopts the double termination convergence criterion [25], that is, continuous unimproved maximum genetic algebra method. In the iterative process, if any of the above rules are met, the iteration will be stopped, and the optimal solution of the smart home product layout optimization model will be output to realize the optimal design of the indoor space layout. The real-coded genetic algorithm solution process is shown in Figure 2:

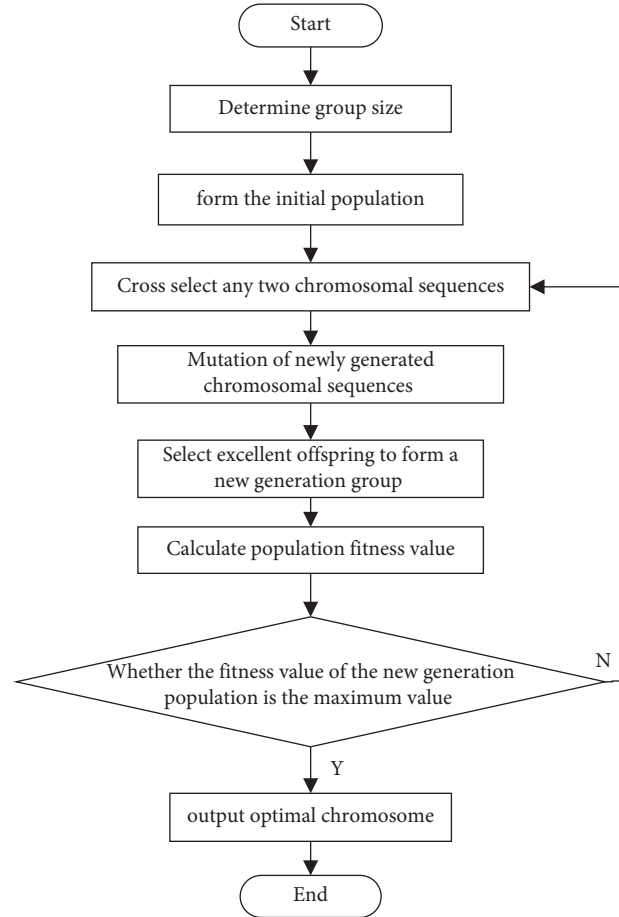


FIGURE 2: The solution flow of the real-coded genetic algorithm.

3. Results and Discussion

3.1. Experimental Parameter Settings. In order to verify the effectiveness of the smart home product layout design method based on the real-number coding genetic algorithm, this paper downloads a large number of (to be counted) furniture models from the 3D model website and categorizes them according to the basic labels, including cabinets, chairs, tables, coffee tables, shelves, dining tables, TVs, etc., and manually consistent size and orientation of each furniture model. The feature extraction simulation of the smart home product layout is carried out, and the visual sampling of the smart home product layout is obtained, as shown in Figure 3.

The parameters of a real-coded genetic algorithm are as follows: the initial population size is 20; Maximum genetic algebra 50; The probability of crossing is 0.8; The window variation probability is 0.04 and 0.8; When the appropriate value of the mutated individual is not better than that of the premutated individual, the probability of replacing the premutated individual is 0.05.

3.2. Experimental Results and Analysis. The performance of different methods for extracting layout features of smart home products is tested, and the comparison results are shown in Table 1.



FIGURE 3: Visual sampling of smart home product layout.

Analysis of the simulation results shows that the method proposed in this paper has a high degree of recognition and a better ability to distinguish the layout features of smart home products, which improves the display and expression capabilities of smart home product layouts.

This paper selects 5 scene instances for comparative analysis, which correspond to the living room, study, bedroom, writing, and dining room, and 5 kinds of home or workplace. The active interaction design method of the smart home dialog agent in reference [4] and the design method of the IoT smart home control system based on 2.4 G

TABLE 1: Feature extraction performance of smart home product layout.

Testing frequency	Feature extraction accuracy (%)	Spatial resolution [dpi]
10	94.6000	4.7684
20	97.3200	4.3474
30	93.9600	4.4000
40	98.2800	4.4947
50	92.6800	4.5579
60	93.3200	4.3158
70	90.9200	4.6737
80	92.2000	4.5474
90	93.8000	4.0526
100	90.7600	4.4000
110	89.4800	4.5158
120	88.5200	4.4000

TABLE 2: Comparison results of smart home product layout optimization design.

Scenes	Number of homes	Pair relationship		
		The method of this paper	Reference [4] method	Reference [5] method
Living room	16	Sofa and TV	Sofa and TV	Sofa and coffee table
		Sofa and coffee table	—	Coffee table and TV
		Coffee table and TV	—	—
Study	20	Table and chair	Coffee table and shelf	Coffee table and stool
		Coffee table and stool	—	Coffee table and shelf
		Coffee table and shelf	—	—
Bedroom	14	Bed and nightstand	Bed and nightstand	Bed and nightstand
		Bed and TV cabinet	—	Bed and TV cabinet
		Dresser and stool	—	—
Office	28	Desk and chair	Work desk and work desk	Desk and chair
		Work desk and work desk	—	—
Dining room	63	Dining table and chairs	—	Dining table and chairs

and zigzag in reference [5] are used as experimental comparison methods, the optimization effect of smart home product layout design is tested, and the comparison results of smart home product layout optimization design are shown in Table 2.

From the analysis of Table 2, it can be seen that due to the problem of constraint conflict, the solution space of the problem is complicated, which makes it difficult for the reference method to obtain the optimal solution in a limited time. The complexity of the furniture arrangement problem is not only related to the number of furniture but also related to the logarithm of the paired relationship between the furniture. The resulting problem is more complicated. Therefore, with more furniture, more crowded rooms, and more complex pairing relationships situation, it will be more difficult to solve. Therefore, the real-number coding genetic algorithm proposed in this paper decomposes the constraint conflict of the paired relationship so that the paired relationship between the furniture has a clear hierarchy. Local optimization and then global optimization can decompose complex problems into simple ones. Therefore, in the optimization design of the home product layout, the paired quantity of the home is accurately classified.

Taking the smart home product layout time as an indicator, test the smart home product layout design efficiency of this method, reference [4] method, and reference [5]

method. During the test, taking the visual sampling of the smart home product layout shown in Figure 3 as the standard, three methods are used to test the smart home product layout of a new room. The test is divided into five groups. The rooms in each group are different in size. The specific test results are shown in Figure 4.

From the data in Figure 4, it can be seen that compared with the method in reference [4] and the method in reference [5], the method in this paper takes a shorter time to design the layout of smart home products and the average time is 1 min, because the method in this paper is used in smart home products. Based on the layout optimization model, the layout design of smart home products is realized by combining the idea of constraint solving and a step-by-step refinement strategy, which shortens the time required for the layout design of smart home products and improves the efficiency of layout design of smart home products.

In order to further verify the effectiveness of this method, this method, reference [4] method, and reference [5] method are used to plan the space of smart home products, and the rationality of layout planning of different methods is compared. During the test, take the average of 200 users' satisfaction with the layout of five groups of different rooms as a result of statistics, and the test results are shown in Figure 5.

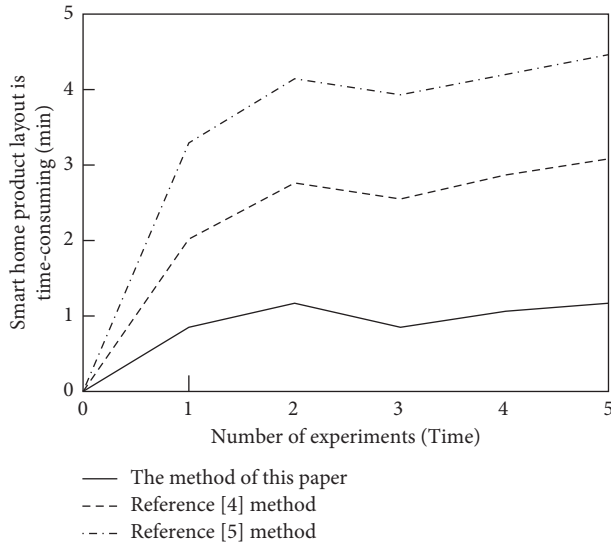


FIGURE 4: Smart home product layout design efficiency test results.

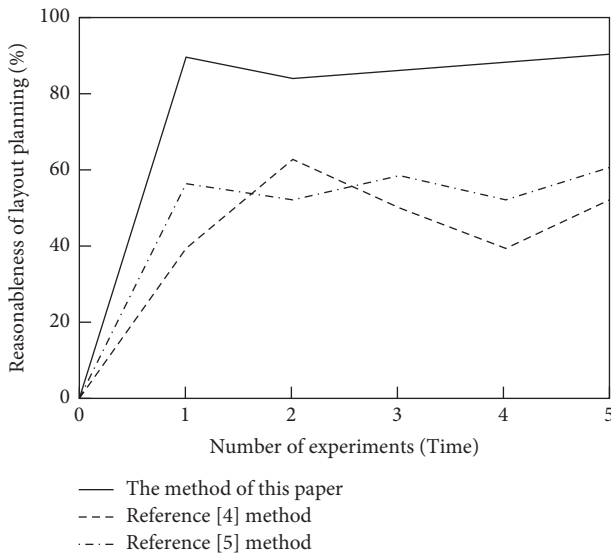


FIGURE 5: Layout planning rationality test results.

Analysis of Figure 5 shows that the rationality of the spatial layout of smart home products after the method in this paper is more than 80%, and the average is 89%, which is much higher than that of the method in reference [4] and the method in reference [5] after the layout planning. The rationality of the space layout of smart home products. The method in this paper establishes a smart home product layout optimization model on the basis of extracting the characteristics of smart home products and uses a real-number coding genetic algorithm to solve it, which improves the rationality of layout planning.

4. Conclusions

There are many kinds of smart home products which are prone to constraint conflicts in the layout process. This paper presents a real-coded genetic algorithm-based smart home

product layout design method. This paper establishes a hierarchical tree based on the prior pair relationship between furniture to degrade the constraint conflicts and builds a smart home based on the two constraints of the total area and individual area of the smart home product space layout; product layout optimization model, which also reduces the complexity of the search problem and improves the speed of the search. In addition, in order to further improve the solution efficiency, this paper introduces the real-number coding genetic algorithm to optimize the solution. The real-coded genetic algorithm has a faster convergence rate, and since it starts searching from multiple initial values at the same time, it also improves the quality of the search. Finally, the experiment is carried out with five scenarios. The experimental results show that the average time of smart home product layout design is 1 min, and the average value of layout planning rationality is 89%. The results show the effectiveness of this method. The scene after the application of this method meets the needs of functionality and aesthetics.

Data Availability

The dataset can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] J. Lu, C. Ou, C. Liao, Z. Zhang, K. Chen, and X. Liao, "Formal modelling of a sheet metal smart manufacturing system by using Petri nets and first-order predicate logic," *Journal of Intelligent Manufacturing*, vol. 32, no. 5, pp. 1043–1063, 2021.
- [2] R. Tang and Y. Inoue, "Services on platform ecosystems in the smart home 2.0 era: elements influencing consumers' value perception for smart home products," *Sensors*, vol. 21, no. 21, p. 7391, 2021.
- [3] D. Pal, C. Arpnikanondt, S. Funilkul, and M. A. Razaque, "Analyzing the adoption and diffusion of voice-enabled smart-home systems: empirical evidence from Thailand," *Universal Access in the Information Society*, vol. 20, no. 17, pp. 1–19, 2020.
- [4] E. Z. Wang, X. Yuan, and S. Y. Li, "Proactive interaction design of conversational agent for smart homes," *Journal of Graphics*, vol. 41, no. 4, pp. 658–666, 2020.
- [5] M. Yin, F. H. Li, and H. Y. Hou, "Design of smart home control system of IOT based on 2.4g and zigbee," *Automation & Instrumentation*, vol. 35, no. 5, pp. 21–25, 2020.
- [6] M. Batool, A. Kim, and K. Kim, "Telemonitoring of daily activity using accelerometer and gyroscope in smart home environments," *Journal of Electrical Engineering & Technology*, vol. 15, no. 6, pp. 2801–2809, 2020.
- [7] M. A. Khan, S. Abbas, and A. R. Rehman, "A machine learning approach for blockchain-based smart home networks security," *IEEE Network*, vol. 35, no. 3, pp. 223–229, 2021.
- [8] H. Zhu, Z. Li, X. Wang, and L. Chen, "Multi-party quantum key agreement protocol for smart home environment,"

- International Journal of Theoretical Physics*, vol. 60, no. 10, pp. 3948–3960, 2021.
- [9] A. B. Fakhri, S. K. Gharghan, and S. L. Mohammed, “Power reduction based on sleep/wake scheme in wireless sensor network for patients vital sign monitoring system,” *Technology Reports of Kansai University*, vol. 62, no. 4, pp. 1669–1685, 2020.
- [10] L. Zhang, “Effective integration of ZigBee communication technology and Internet of things technology,” *Journal of Interconnection Networks*, vol. 22, no. 3, Article ID 2145011, 2022.
- [11] A. H. M. Alkawgani, A. R. H. Alhawari, A. T. Hindi, and A. Y. A. Ashwal, “Hybrid image steganography method using Lempel Ziv Welch and genetic algorithms for hiding confidential data,” *Multidimensional Systems and Signal Processing*, vol. 33, no. 2, pp. 561–578, 2022.
- [12] A. Yu and S. Yang, “Research on web server cluster load balancing algorithm in web education system,” *The Journal of Supercomputing*, vol. 76, no. 5, pp. 1–10, 2020.
- [13] L. Papageorgiou, H. Alkenaris, M. Zervou, D. Vlachakis, and E. Eliopoulos, “Pos0350 epione application: an integrated genotype analysis web server for clinical genomics in systemic lupus erythematosus (sle),” *Annals of the Rheumatic Diseases*, vol. 80, no. Suppl 1, p. 405, 2021.
- [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. 1–15, 2020.
- [15] B. N. Sa Nd Aker, B. Kleven, and A. R. Wang, “Structural typologies and the architectural space—studies of the relationship between structure and space by application of structural types to multistory buildings,” *Architecture, Structures and Construction*, vol. 2, no. 1, pp. 199–221, 2022.
- [16] F. C. Eroglu, I. Inanc, B. Sabuncuoglu, E. Sen, and B. Kazanc, “The relationship between histopathological features and surgical success in primary acquired punctal stenosis,” *European Journal of Ophthalmology*, vol. 32, no. 3, pp. 1457–1463, 2022.
- [17] A. Vincent, A. Kumar, and P. Upadhyay, “Effect of red-edge region in fuzzy classification: a case study of sunflower crop,” *Journal of the Indian Society of Remote Sensing*, vol. 48, no. 10, pp. 645–657, 2020.
- [18] H. Wang, L. Zhu, G. Li, M. Zuo, X. Ma, and J. Wang, “Perfusion parameters of intravoxel incoherent motion based on tumor edge region of interest in cervical cancer: evaluation of differentiation and correlation with dynamic contrast-enhanced mri,” *Acta Radiologica*, vol. 61, no. 8, pp. 1087–1095, 2020.
- [19] S. Nezhad, M. Zoj, and A. Ghorbanian, “A fast non-iterative method for the object to image space best scanline determination of spaceborne linear array pushbroom images,” *Advances in Space Research*, vol. 68, no. 9, pp. 3584–3593, 2021.
- [20] M. Ismail, M. Sale, Y. Yu et al., “Development of a genetic algorithm and NONMEM workbench for automating and improving population pharmacokinetic/pharmacodynamic model selection,” *Journal of Pharmacokinetics and Pharmacodynamics*, vol. 49, no. 2, pp. 243–256, 2022.
- [21] R. E. Precup, E. L. Hedrea, R. C. Roman, and E. M. Petriu, “Experiment-based approach to teach optimization techniques,” *IEEE Transactions on Education*, vol. 64, no. 2, pp. 1–7, 2020.
- [22] A. S. Akopov, L. A. Beklaryan, and A. L. Beklaryan, “Cluster-based optimization of an evacuation process using a parallel bi-objective real-coded genetic algorithm,” *Cybernetics and Information Technologies*, vol. 20, no. 3, pp. 45–63, 2020.
- [23] M. Ramanauskas, D. Eok, J. Ilinskas, V. Starikovicius, and A. Kaceniauskas, “Global optimization of grillage-type foundations using a distributed genetic algorithm,” *Journal of Global Optimization*, vol. 77, no. 15, pp. 157–173, 2020.
- [24] A. C. Pandey, A. Kulhari, and D. S. Shukla, “Enhancing sentiment analysis using roulette wheel selection based cuckoo search clustering method,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 1, pp. 1–29, 2021.
- [25] X. Y. Xue, S. G. Zhao, and G. Cheng, “Simulation of database lost data recovery and reconstruction based on genetic optimization,” *Computer Simulation*, vol. 38, no. 11, pp. 485–489, 2021.

Research Article

Visual Network Analysis Based on Stereo Vision and Feature Matching Algorithm

Lili Shi 

Qingdao Huanghai College, Qingdao, Shandong 266000, China

Correspondence should be addressed to Lili Shi; shill@qdhhc.edu.cn

Received 15 April 2022; Revised 28 May 2022; Accepted 31 May 2022; Published 28 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Lili Shi. 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.

Functions such as Internet browsing and online shopping have a great impact on people's lives. The footprints of people browsing various information and news on the web page are also increasing year by year. More and more people begin to pay attention to the visual communication in web design. Network information is widely loved by people because of its convenience, quickness, and simplicity. In order to study the visual problem in network information, this paper proposes to use a feature matching algorithm to study the visual information transmission in web design. Using stereo vision and feature matching algorithm, the target recognition function in the visual communication of web design is realized. The content of web design is defined from the perspective of visual beauty and overall harmony. By extracting the feature points in the data, the two-dimensional vision is transformed into a three-dimensional vision. It not only can accurately extract feature points from data but also can convert 2D vision to 3D vision. Finally, in order to optimize the feature matching speed of web design images, the epipolar constraint algorithm is used to optimize the feature matching function. The experimental results show that the content of this paper can truly show the virtual effect in web design and can intuitively upload and transmit simple and rich information content. This paper not only meets the aesthetic needs of the public for web design but also improves the problem of fuzzy information in the process of visual information transmission.

1. Introduction

Web design is the effect of information transmission through the combination of network technology and computer technology with visual elements [1]. When browsing websites and various web pages, people usually judge and evaluate the beauty and visual experience of web pages. If network information wants to become the main way of multimedia information transmission, it is necessary to optimize and update the field of Web Design [2]. Focus on the effect of visual communication and people's experience. Make visual communication and web design meet the needs of the masses. Therefore, we need to analyze various visual elements in website construction and web page design to determine the impact of visual effects on people's physiological needs [3]. Professional website designers also need to investigate design schemes with high public feedback and integrate visual elements into web browsing. By analyzing

and evaluating the connection, location, and feedback of visual senses, researchers explored the human visual process. This paper mainly analyzes the internal structure of the eye in the process of human vision [4]. When we observe objects and browse the web, we only focus on the objects we are interested in. Focus on the focus first and then pay attention to the surrounding environmental factors [5]. Therefore, the browsing process is a dynamic sequential structure, and there is order and direction. We can judge the current psychological activities and interest characteristics of users according to the sense of direction of visual browsing [6].

With the gradual improvement of the quality of people's daily life, there are richer requirements in aesthetic needs. Nowadays, web design has shifted from functionality to aesthetics [7]. It takes network information as data and a network platform as a carrier to transmit all kinds of news and newsletters to people. This convenient and fast way has been highly praised by everyone. It is precisely because

people's aesthetic needs are gradually increasing that people's visual feelings need to be considered in web design [8]. If we need to beautify and improve the visual function, we should first ensure the simplicity of information content in web page design and serve the form of visual communication. Second, the information content should be displayed in the visual process to ensure a natural and smooth browsing feeling [9]. At the same time, the quantity of network information is controlled, and the overall layout is planned uniformly. Finally, in web design, we need to ensure that people can focus on the information content, and cannot miss the key content because of the surrounding environment or useless information on the web page. In order to meet the above problems in the visual communication of web design, we use a feature matching algorithm for research and analysis [10].

This paper is mainly divided into three parts. The first part briefly analyzes the visual information transmission function of web design and explores the application status of the feature matching algorithm in various countries. The second part first analyzes various elements in web design and users' needs for visual elements. Adopt stereo vision to transform people's traditional perception of web design and develop from 2D design to 3D design. Second, the point target recognition function of the feature matching algorithm on visual information is studied. Finally, in order to further improve the visual information transmission effect, we use the epipolar constraint algorithm to optimize it. Explore the effect of web page design optimized by feature matching algorithm. The third part analyzes the results of the research on visual communication of web design based on feature matching optimization algorithm.

2. Related Work

At present, web designers have a poor understanding of the role and importance of visual communication information and have not formed a deep understanding [11]. With the continuous development of information technology, they found that the use of web visual elements is not much in their application and research. Therefore, in the face of the slow development of visual information communication, we analyze the problems of web design at this stage. First, the visual expression function of web design is relatively single, and the amount of data of news information is large, which makes it difficult to typeset and classify. The accumulation of website information is a difficult problem for most designers. This not only limits the network communication function but also causes great trouble to visitors. The visual expression of web pages is also relatively single, which cannot attract users' interest in form. Second, although web design needs to accurately express information, most websites are not beautiful enough. The aesthetics of web pages is the core content of visual information transmission. We need to improve the typesetting efficiency of web information and reasonably match the color structure [12]. Finally, in the problem analysis, we found that the interaction between websites and web pages and users is too poor to help users obtain information quickly. People are easily affected by

visual elements, resulting in distraction. Facing the above problems, we explore the function and function of the feature matching algorithm.

Information technology in the United States has developed rapidly, and they have made many achievements in the research of image feature detection [13]. The process of feature point detection and matching in computer vision tasks is a complex structure. Compared with the traditional image system, the image extraction system using a feature matching algorithm can label the scene information and improve the accuracy of feature detection.

The research cost invested by Germany in the field of UAV is large [14]. In order to meet the concept of low-altitude photography, they use UAV images to complete the photography link. However, in practical application, there are some problems with the image data taken by UAV, such as small pixels, more interfering objects, small photographing range, and so on. Therefore, they use a feature matching algorithm to optimize UAV photography. It improves the situation of image blur and many interference factors.

China's concept of green environmental protection is an important link to national sustainable development [15]. In the research of forestry information construction, there are high requirements for the accuracy of data resources and feedback efficiency. In the actual data acquisition, most of the information is easily disturbed by miscellaneous data. Therefore, the traditional forestry information construction still depends on the manual mode. With the development of information technology, researchers use feature matching algorithms to improve the time-consuming links in forestry work. It not only improves the accuracy of data information but also provides optimization and guarantees the processing efficiency of construction work. According to the above research on the development status of the feature matching algorithm in various countries, we have applied this algorithm to the analysis of visual information communication in web design and achieved good results.

3. Methodology

3.1. Research on Stereo Vision Based on Feature Matching Algorithm to Realize Point Target Recognition in Web Page Design. Information communication in web design is an important part of media communication. With the rapid development of network technology, the forms of visual communication have also changed. Technology is the basis of realizing the visual effect of web pages. Designers should take the initiative to master various existing network technology laws, pay attention to the close combination of technology and art, make full use of the advantages of technology, realize artistic imagination, and achieve ideal visual effects. Excellent layout design expresses harmony and beauty through the spatial combination of words and graphics. It can optimize reading, accurately transmit information, and make the website have an affinity. In web design, different colors should be combined and matched according to the principles of harmony, balance, and emphasis. In the network age, website and web page design

need to meet the premise of functionality and ensure vision and aesthetics. At present, the focus of web design is still biased towards the literal mode, and the designer's technology and thinking are also limited. The visual expression of web pages is relatively single, not beautiful and advanced enough. In order to improve the conversion from the original two-dimensional image target to a three-dimensional image, we use the point target recognition technology of stereo vision to optimize the web design. First, the 3D information in the data is obtained; the 3D visual model is established; and the boundary of the data points is judged and analyzed. The acquisition of point data is mainly recorded by laser scanning. In order to better store and match data, we use a ring network for the construction of three-dimensional data points. First, the target position of the data is defined as the origin of the coordinate system, and the vertical parameters are injected into the coordinate system. Through the above operations, we can obtain the optimal position distance of visual information points. In order to meet the coverage and accuracy of three-dimensional direction, we express the optimal distance as

$$D_{\text{best}} = \min \left[\sqrt{(x_i - x_o)^2 + (y_i - y_o)^2 + (z_i - z_o)^2} \right], \quad (1)$$

where i and o represent the target point coordinates and the simulation coordinate system, respectively. The minimum distance in all positions can be obtained by calculating the three directions corresponding to the feature points. Because the position of 3D target points will involve attitude problems, we need to model the target point data to test the specific position of the target in the coordinate system. The center relationship of the datum line can be expressed as follows:

$$\begin{aligned} X_i &= (a + D_{\text{best}})\cos(\alpha_i)\cos(\beta_i), \\ Y_i &= (b + D_{\text{best}})\sin(\beta_i), \\ Z_i &= (c + D_{\text{best}})\sin(\alpha_i)\cos(\beta_i), \end{aligned} \quad (2)$$

where X_i , Y_i , and Z_i is the center position parameter of the datum line, which can calculate the correction range of the optimal distance. α_i and β_i represent the offset value of the corresponding target point. Through the test of the offset value, we can judge whether there is a difference in the angle.

The essential concept of visual stereo is to reorganize the three-dimensional coefficients of the target to be measured by analyzing the three-dimensional structure of the feature point target. In the actual recognition process, because the target has an external structure, the model of the sample data can be used as a reference to judge the position of the target point. In the research on the influencing factors of the stereo vision effect, the most obvious influencing parameter is the angle of the actual position. The distributed recognition effect can be obtained by projecting the angles of different samples, which can improve the success rate of sample feature target recognition. The specific operation flow is shown in Figure 1.

Feature points are those points that are analyzed by the algorithm and contain rich local information. The so-called "scale invariance" of feature points refers to their unified properties that can be recognized in different pictures. As can be seen from Figure 1, first, the position of the detection target is modeled to form different types of feature point sets. The target is projected according to the feature point set. Second, in the projection interval of each feature point, the appropriate proportion is selected for calculation. The final data obtained by calculation is the selected proportional value. After selecting the target sample test area, traverse and filter. Finally, target recognition is carried out, and the number of features is obtained from the target set of multiple structures. The false target is eliminated by similarity calculation to obtain the final stereo vision effect.

In practice, the target interval to be selected can be represented by any parameter. When the laser scanning is projected to the target, the distance parameter between samples can be obtained. Because of stereo vision, 3D style changes from 2D data. Therefore, if there is a given target at the corresponding projection angle, the stacking data along the fuzzy position of the target can obtain the specific position. If the projection distance is too large, the value calculated by traversal is obtained according to a certain proportion. The data of target points is in three-dimensional mode. Before completing the construction of stereo vision, the angle of two-dimensional to three-dimensional mapping needs to be selected. It is known that the laser direction is a fixed parameter and the direction is the real distance of the projection. Therefore, the laser direction is selected as the projection direction when extracting the feature of the target point set. In order to optimize the two-dimensional style of web design, the data of sample parameters in feature points are complex. These eigenvalues can automatically form any matching matrix. We calculate the eigenvalues of the matrix to obtain the actual matching relationship between targets. To determine whether the target conforms to the three-dimensional type, the matrix expression is

$$\begin{aligned} A &= (r - 1)^{-1} C^T C = \begin{pmatrix} \text{cov}(x, x) & \text{cov}(x, y) \\ \text{cov}(y, x) & \text{cov}(y, y) \end{pmatrix}, \\ A &= (r - 1)^{-1} C^V C = \begin{pmatrix} \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, y) & \text{cov}(y, z) \end{pmatrix}, \\ A &= (2r - 1) C^V C^T = \begin{pmatrix} \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, y) & \text{cov}(z, z) \end{pmatrix}. \end{aligned} \quad (3)$$

In (3), the matrix has a covariance form; this structure is symmetrical; and the calculation results can be obtained by eigenvector decomposition. For different characteristic functions, the actual positions of test points are similar. In order to distinguish effectively, we use a similar function calculation:

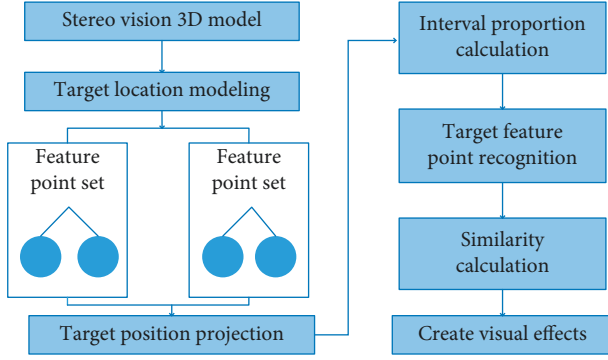


FIGURE 1: Flow chart of stereo vision target recognition.

$$r(fp, fQ) = \frac{n \sum fpfQ - \sum fp \sum fQ}{\sqrt{n \sum f^2 p - (\sum fp)^2}}, \quad (4)$$

$$r(fp, fQ) = \sqrt{\frac{n \sum f^2 Q - (\sum fQ)^2}{n \sum f^2 p - (\sum fp)^2}},$$

where p and Q represent two test target points, and fp and fQ are the eigenvalue set range corresponding to the two targets. n represents the number of calculated eigenvectors. In order to improve the accuracy of target recognition, we bring stereo vision into the feature matching algorithm of boundary value calculation and filter the position according to the value range in the feature matching process. Only the data that conforms to the defined range of the target can be matched, which greatly improves the information accuracy of the visual effect. The overall beauty of the web page design optimized by stereo vision has been improved. We investigated this and formed a comparative evaluation, as shown in Figure 2.

As can be seen from Figure 2, we selected loyal customers who have used a website for a long time as variable parameters and the evaluation results as control variables. The stereo vision effect optimized by the feature matching algorithm has achieved good evaluation feedback in web page design. Compared with the visual communication effect of traditional web design, the research content of this paper is effective.

3.2. Research on Visual Information Communication of Web Design Based on Feature Matching Algorithm Optimization. Visual communication design plays a very important role in improving the competitiveness of Chinese enterprises. The development speed is very rapid in the intersection and blending with many disciplines. Due to the differences in solving problems and different ways of solving problems, different knowledge structures are required, so new majors are bound to appear and develop. Once the design is simplified by computers and other modern equipment, most of the designer's time will not be spent on the completion work. The formation of ideas plays an important role. The design and function of web pages can help people get information, and the importance of visual communication cannot be ignored. Visual

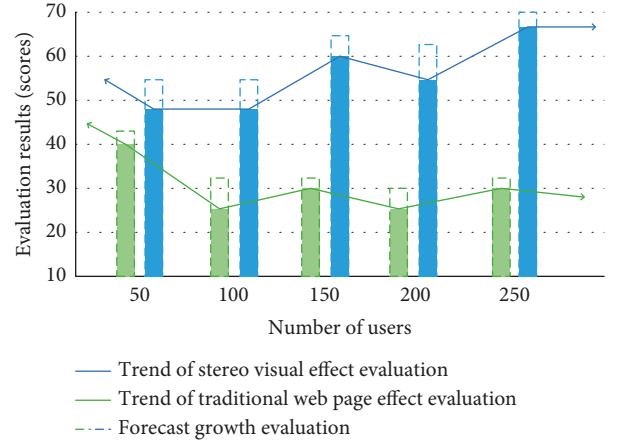


FIGURE 2: Comparison of evaluation before and after optimization.

information transmission should take the transmission speed of information as the primary research content and improve the accuracy of the information in practical optimization. Finally, combined with the aesthetic trend of the masses, the web page design is optimized and improved. In the environment of the big data era, web pages will produce news information with a huge amount of data every day. Concise typesetting and layout are very important links in web design. In order to ensure that users can easily obtain the required content, we consider simplifying the processing of the information. However, in the process of information simplification, it is easy to reduce the accuracy. In this regard, our research needs to ensure the accuracy of information and meet the visual effect of web design. In order to improve the matching speed and accuracy of information content, we optimize the feature matching algorithm. A matching algorithm RANSAC algorithm based on epipolar constraint is proposed, which processes the feature point data in the information and processes the information normally. The original feature matching uses the stable data feature points in the information to compress the amount of information processing to achieve the effect of search and matching. This search mode has less computation and is easier to complete the specified task. However, in the big data environment, such methods can no longer meet the current needs of people. Therefore, we use the epipolar constraint matching algorithm to optimize and summarize the information feature points into the same epipolar, which greatly reduces the scope of matching and search. The feature point basis vector matches the feature points to explore the relationship between them. The calculation results can directly affect the accuracy of the matching information. We normalize the information data and convert the original information to reduce the impact of interference data. RANSAC parallel algorithm is used to improve the speed of calculation. First, change the location of the matching information and standardize the location. Assuming that the parameters of any feature point are a function value, the processed information points are

$$\begin{aligned} M_l &= T_l m_l, M_r = T_r m_r, \\ M_l &= T_l m_l, M_r = T_r m_r, \end{aligned} \quad (5)$$

in which

$$T_l = \begin{bmatrix} k_0 - k_m u_0 \\ 0k - k_m v_0 \\ 0 \dots 0 \dots 1 \end{bmatrix},$$

$$u_0 = \frac{1}{N} \sum_{i=1}^N u_i, \quad (6)$$

$$v_0 = \frac{1}{N} \sum_{i=1}^N v_i,$$

$$km = \frac{\sqrt{2}}{\sqrt{(1/N) \sum_{i=1}^N [(u_i - u_0)^2 + (v_i - v_0)^2]}}$$

In order to improve the efficiency of information computing, we use parallel computing to obtain matrix results. Record the maximum number of values in each group and the corresponding basic matrix and make statistics by sampling count:

$$m = \frac{\log(1-p)}{\log[1-(1-\lambda)^8]}, \quad (7)$$

where p represents the probability that at least one of m sample points is accurate information and λ represents the error rate of matching. The distance between the data information matching points is calculated according to the set characteristics of the epipolar line, and the calculation results are used to express the deviation variable. The corresponding formula is

$$d_i = d_l + d_r = d(m_{l1}, F^T m_{r2}) + d(m_{l2}, F m_{r1}), \quad (8)$$

where d_i represents the distance from the information point to the polar line. We select a piece of information data in the sample web page as the test point. The positional relationship between sample feature points and epipolar lines is shown in Figure 3.

It can be seen from Figure 3 that the data in the left figure are scattered greatly before eigenvalue matching optimization, while the data after feature point matching optimization are related to the epipolar position. It can be seen that each calculation process of the RANSAC algorithm model is carried out independently. Selecting a large number of initial data of information samples can also improve the calculation speed of the algorithm.

4. Result Analysis and Discussion

4.1. Analysis of Research Results of Point Target Recognition Technology Based on Feature Matching Algorithm. Whether websites and web pages can convey effective information to users is affected by many factors. Among them, the visual communication effect is a very important link. The visual effect of the web page can bring users the impression

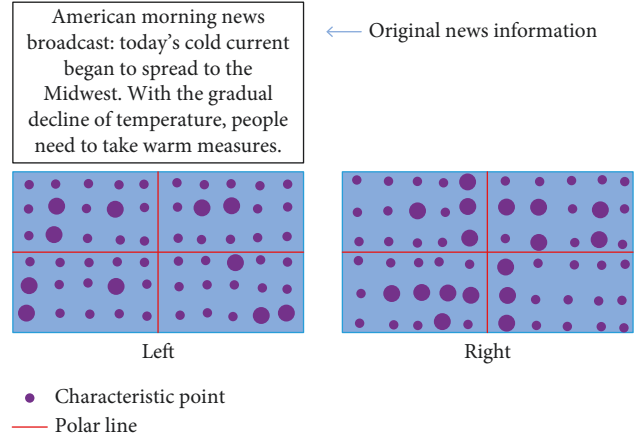


FIGURE 3: Position relationship between sample feature points and epipolar lines.

that meets their aesthetic needs. It is the first environment for people to contact information. Information goes deep into the brain through visual communication functions, and people analyze and interpret information. Therefore, the concept and thinking of web design must comply with the visual information communication effect and visual communication process. The visual information transmission process is controlled by people's eye characteristics. Physiological conditions and eye structure are complex environments. We understand and absorb information through visual target capture. In web design, designers should pay attention to the visual effect of web pages. The traditional two-dimensional space effect can no longer meet the needs of modern society. In order to improve the visual information transmission effect of web pages, this paper uses a feature matching algorithm to upgrade the two-dimensional effect. The design method in this paper has a stronger sense of space in aesthetics and can bring more intuitive feelings to readers. In order to make a more accurate model, research has improved the visual information transmission effect of web pages. In the experimental process, laser scanning is used to process the point target data in the area. The test object is aimed at the image information inside the web page. Detect and recognize different images and types. First, the target point data is obtained by laser scanning. Then, the filtering operation is performed to obtain the image information marked by the feature points. Because the sample data to be tested is filtered for a long time, the sample points in line with visual aesthetics will be retained. Remove excess astigmatism, background, and other interference data and reanalyze the test image from various angles. Finally, the marked target is consistent. In order to test the recognition effect of different image information in web design. We control the experimental conditions and set two modes: multitarget coincidence and no occlusion. Explore the integrity of stereo vision effect on image information expression, as shown in Figure 4.

It can be seen from Figure 4 that the stereo vision studied by using the feature matching algorithm has the highest effect on the integrity of image information representation. The traditional visual information transmission function

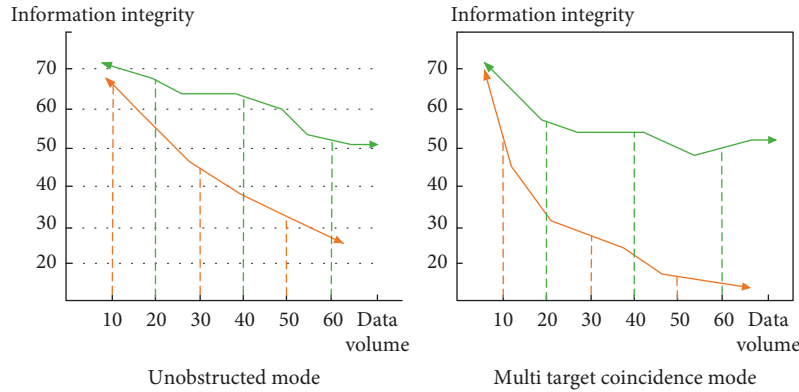


FIGURE 4: Comparison of integrity between stereo visual information expression and traditional visual expression.

cannot effectively obtain and recognize the information in the face of multitarget coincidence. Therefore, the three-dimensional effect of the feature matching algorithm in stereo vision can restore image information in a complex environment, which contributes to the beauty and accuracy of web design.

4.2. Analysis of Research Results of Visual Information Communication in Web Design Based on Feature Matching Algorithm Optimization. In order to improve the accuracy and efficiency of information acquisition in web design, the function of visual information transmission should also be optimized. Designers need to follow the visual and sensory characteristics of the masses, start from the information and data typesetting, and realize the aesthetic needs and interactive functions of the masses. In the face of more big data information, it is necessary to simplify the information data. Prevent complex environment and messy information from affecting users' attention. In the visual information transmission of web design, in order to improve the information accuracy and matching effect in the big data environment, we optimize the feature matching algorithm. RANSAC algorithm based on epipolar constraint is proposed to control the distance of feature points. In order to verify that the experimental algorithm improves the accuracy of information acquisition, we use the internal information of a web page in big data as the experimental data. Take the number of feature points of the target to be measured as the variation and explore the information accuracy before and after the optimization of the feature matching algorithm, as shown in Figure 5.

As can be seen from Figure 5, the experimental results of the two algorithms are very different. Due to the huge amount of data information, the traditional feature matching algorithm suffers from data loss when identifying information targets. This leads to a decline in the internal accuracy of information as the amount of information increases. The optimized feature matching algorithm not only can ensure the accuracy of information data but also can meet the visual aesthetics of the masses. Finally, this paper puts forward some strategies for visual information communication in web design. First, designers should pay

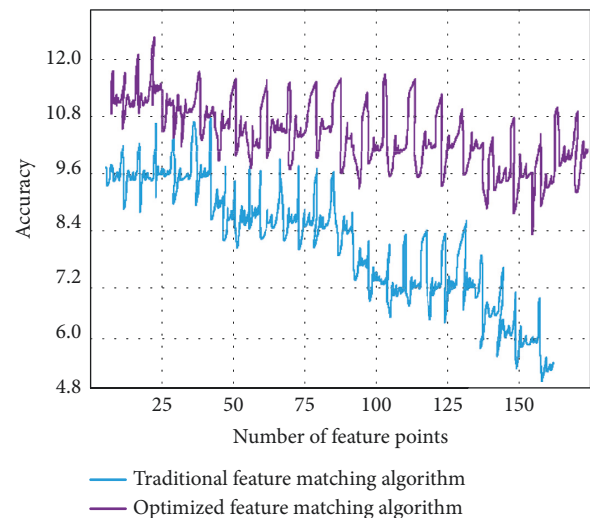


FIGURE 5: Comparison of information accuracy before and after feature matching algorithm optimization.

attention to the main content and aesthetic needs of information on web pages. While ensuring users' perception of visual senses, highlight the key content of information. Highlight their own advantages and characteristics in website design and web page layout and bring a convenient and fast user experience to users. Second, the rational use of multimedia information technology not only can add picture elements to web design but also can add dynamic elements such as video and audio. This multimode design is more in line with people's visual experience than traditional text information. Finally, the interaction effect with users is added to web design. Let users change from passive acceptance of information to active acquisition.

5. Conclusion

Web browsing is the main way to obtain information in the era of big data. As a form of visual communication, web design can combine text, pictures, audio, and other elements and become an interactive tool to meet multidimensional feelings. Traditional web design can only meet the expression of information. With people's longing for a better life,

this form of web design does not meet the needs of contemporary users. Based on the above situation, this paper uses the feature matching optimization algorithm to study the visual information transmission mode of web design. This paper analyzes and explores the situation of single web page arrangement, poor information transmission effect, and poor interactivity. First, the three-dimensional visual effect is used to obtain the target information on the web page. The main contents of the information are marked, and the matching function is optimized by a feature matching algorithm. Second, focus on the problem of recognition accuracy, highlight the target information from a variety of web pages, and realize the simple and fast function of visual communication. Finally, we find that the traditional feature recognition algorithm is easy to cause information loss when matching information in the big data environment. In this paper, the RANSAC algorithm is used to optimize feature matching and apply it to the visual information communication function. The experimental results show that the feature matching algorithm can accurately obtain the web page information and replan the stereo vision. Compared with the traditional two-dimensional mode, it has been greatly improved in design aesthetics. The optimized feature matching algorithm can also ensure the accuracy of the information in the case of too much web page data and improve users' experience of web pages. However, the research still has some limitations. This paper does not discuss the visual performance of the feature recognition algorithm in different web page features. Therefore, there is a problem of weak applicability. This needs further elaboration and analysis in future discussions.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by Qingdao Huanghai College.

References

- [1] Q. Zhang, "Research on text graphics in visual communication design," *Journal of Hubei open vocational college*, vol. 34, no. 13, pp. 196–198, 2021.
- [2] Na Wu, "Color matching of visual communication in web design," *Popular colors*, vol. 07, pp. 34–35, 2021.
- [3] Z. Zhang, "Research on the application of visual communication in web design," *Art education research*, vol. 10, pp. 76–77, 2021.
- [4] Z. Cao, "Expression and application of visual elements in web design," *China newspaper*, vol. 06, pp. 64–65, 2021.
- [5] J. Zhao, "Application of information design in web visual communication," *Modern industrial economy and informatization*, vol. 11, no. 02, pp. 128–129 + 132, 2021.
- [6] Li Hui, "Research on text application based on visual communication design," *Journal of Jingdezhen University*, vol. 36, no. 01, pp. 114–118, 2021.
- [7] M. Sun, "Research on the application of visual elements in Web Design in the new media era," *Editing*, vol. 2021, no. 01, pp. 39–40, 2021.
- [8] F. Zhang and M. Li, "Plane visual communication design of user experience effect based on computer simulation," *Journal of Physics: Conference Series*, vol. 2143, no. 1, Article ID 012026, 2021.
- [9] H. Guo, L. Xuan, C. Hao, F. Kateb, and A. A. Y. Alzyoud, "Research on the mathematical model construction of art and ideology in visual communication design teaching," *Applied Mathematics and Nonlinear Sciences*, vol. 6, no. 2, pp. 517–524, 2021.
- [10] H. Qin, J. Xiong, and M. Xu, "Research on the unity of visual communication and information transmission in web design," *Tomorrow fashion*, vol. 04, pp. 63–65, 2020.
- [11] H. Sun and T. Zheng, "Expression form of web design in visual communication design," *Western leather*, vol. 42, no. 03, p. 92, 2020.
- [12] Z. Zhang, Z. Zhang, B. Zhu et al., "A fast algorithm for feature matching and target location based on Shi Tomasi and improved LBP," *Journal of Jilin University (Science Edition)*, vol. 59, no. 05, pp. 1171–1178, 2021.
- [13] Y. Li, W. Zheng, X. Liu, Y. Mou, L. Yin, and B. Yang, "Research and improvement of feature detection algorithm based on FAST," *Rendiconti Lincei-Scienze Fisiche e Naturali*, vol. 32, no. 4, pp. 775–789, 2021.
- [14] K. Ye, Y. he, J. Chi, X. Zhu, F. Yang, and X. Rui, "Research on visual SLAM algorithm based on optimal feature point extraction," *Industrial control computer*, vol. 34, no. 07, pp. 76–79, 2021.
- [15] G. Ren, P. Liu, and Z. He, "Improved matching algorithm based on lidar feature extraction," *Journal of Shaanxi University of Science & Technology*, vol. 39, no. 03, pp. 138–144, 2021.

Research Article

On Trees with Greatest F – Invariant Using Edge Swapping Operations

Wenhu Wang,^{1,2,3} Adnan Aslam ,⁴ Muhammad Ahsan Binyamin ,⁵ Salma Kanwal ,⁶ and Iqra Irshad⁴

¹School of Software, Pingdingshan University, Pingdingshan, Henan 467000, China

²International Joint Laboratory for Multidimensional Topology and Carcinogenic Characteristics Analysis of Atmospheric Particulate Matter PM2.5, Pingdingshan, Henan 467000, China

³College of Computing and Information Technologies, National University, Manila PH1008, Philippines

⁴University of Engineering and Technology, Lahore(RCET), Pakistan

⁵Department of Mathematics, GC University Faisalabad, Pakistan

⁶Department of Mathematics, Lahore College for Women University, Lahore, Pakistan

Correspondence should be addressed to Muhammad Ahsan Binyamin; ahsanbanyamin@gmail.com and Salma Kanwal; salma.kanwal055@gmail.com

Received 2 May 2022; Accepted 11 June 2022; Published 28 June 2022

Academic Editor: Gohar Ali

Copyright © 2022 Wenhu 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 F -index of a graph Q is defined as $F(Q) = \sum_{t \in V(Q)} (d_t)^3$. In this paper, we use edge swapping transformations to find the extremal value of the F -index among the class of trees with given order, pendent vertices, and diameter. We determine the trees with given order, pendent vertices, and diameter having the greatest F -index value. Also, the first five maximum values of F index among the class of trees with given diameter are determined.

1. Introduction

Mathematical chemistry is providing effective and time-saving methods for evaluating the properties of chemical compounds without having to go through tedious laboratory experimentations. Topological indices are function maps that identify key computational and topological aspects of a structure and evaluate chemical compound properties without using quantum mechanics as final production [1]. The total π -electron energy (E) [2] of a molecule was found to be related to its thermodynamic stability that depends on the structure of a molecule that is its topology. Relationship between (E) and topology of a molecule was determined by its graphical structure [3]. Comparison was made between

the original vertex degree-based indices and lately defined edge degree-dependent indices (termed as reformulated Zagreb indices), while relating the two versions of indices, the relation existing between the graph and its line graph was utilized. Yang et.al. [4] brought into consideration to researchers the relation between the subtree number index and the Wiener index in the class of spiro chains and polyphenyl hexagonal chains.

In this paper, we consider only simple finite and connected graphs. In a graph Q , we denote its vertex set and edge set by $V(Q)$ and $E(Q)$, respectively. Let $d_Q(p)$ denotes the degree of a vertex p . The distance between two vertices $p, t \in V(Q)$ is denoted by $d(p, t)$ and is defined as the length of the shortest path joining them. For more undefined terminologies related

to graph theory, we refer [5]. The first topological index were proposed by Wiener [6] (namely, the Wiener index), while he was working on the boiling point of paraffin. The Wiener index is denoted by $W(Q)$ and is defined as

$$W(Q) = \sum_{\{p,t\} \subseteq V(Q)} d_Q(p,t). \quad (1)$$

Zagreb indices were introduced by Gutman et al. [2] that depend on degrees of nodes and are defined as

$$\begin{aligned} M_1(Q) &= \sum_{p \in V(Q)} (d_p)^2 \\ M_2(Q) &= \sum_{pt \in E(Q)} d_p d_t. \end{aligned} \quad (2)$$

These terms were recognized to be a measure of the extent of branching of the carbon atom skeleton of the underlying molecule. Later, its additive version was brought into kind attention to researchers in [7], which as expected, revealed more hidden chemical properties of chemical compounds. This index is named as the general sum connectivity index, given as

$$\chi_\alpha(Q) = \sum_{pt \in E(Q)} (d_p + d_t)^\alpha. \quad (3)$$

Furtula *et al.* [8] in 2015 introduced the F -index, also referred as the forgotten topological index, which is defined as

$$F(Q) = \sum_{t \in V(Q)} (d_t)^3. \quad (4)$$

This index is also a measure of branching and has same measure of predictability as that of the first Zagreb index. In case of the acentric factor and entropy, both $M_1(Q)$ and $F(Q)$ have a correlation coefficient greater than 0.95 [8].

Ali et al. [9] put forward the survey of work done on the Randic index for certain values of α . Azari et al. [10] considered the forgotten topological index in detail and determined the bounds of this index in terms of other graphical parameters. They analyzed the relationship of this index with already existing versions of Zagreb indices. Z. Che et al. [11] determined new bounds for the forgotten index in terms of graph irregularity, Zagreb indices, and many other existing graph invariants. Further they characterized the graphs attaining these bounds and proved that these newly attained bounds are sharper than the existing ones. Another version of the forgotten index namely the forgotten co-index was brought into attention by Ghalavand et al. [12]. The authors found bounds for this index and provided an ordering of graphs with respect to this index. Gutman et al. [13] provided a finite ascending sequence of the forgotten index for trees and moreover for graphs having some particular values

of the cyclomatic number γ . Gutman et al. [14] proved two weighted inequalities of real nonnegative sequences and then used them to determine lower bounds of certain degree dependent indices.

The main motivation behind this work is the idea practiced in [15], in which authors introduced some edge swapping operations on graph structures and analyzed the behavior of generalized sum connectivity descriptor. The authors found the decreasing behavior of the descriptor and provided the least five values of this descriptor for trees. Further they also provided the trees that attain these least values. In this work, making use of certain graph transformations that involve the swapping of edges from one node to another and contraction of edges, we have observed the behavior of the F -index. This enabled us to determine the decreasing sequence of values of F -invariant and the corresponding trees attaining these values. Novelty of work lies behind the fact that solving a research problem that is not solved already is always a good addition to the existing literature. Thus, this problem of determining members in a certain family of graph with first, second up to fifth extremal values has become good source of attraction to researchers.

2. F -Invariant under Certain Transformations

In this section, we first observe the increasing or decreasing behavior of F -invariant under certain graph operations involving swapping of edges from one node to another. Our next results show that this descriptor exhibit increasing behavior.

$$\alpha_1 - \text{transform.} \quad (5)$$

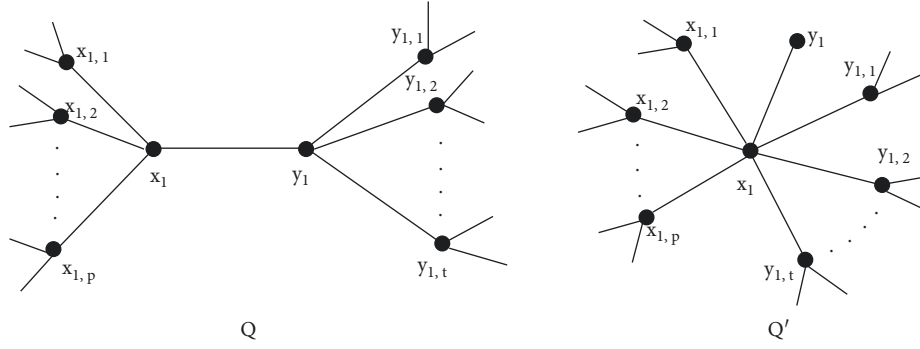
Let Q be the connected tree and $x_1, y_1 \in V(Q)$. For $p \geq 0, t \geq 1$, suppose $N(x_1) = \{y_1, x_{1,1}, x_{1,2}, \dots, x_{1,p}\}$ and $N(y_1) = \{x_1, y_{1,1}, y_{1,2}, \dots, y_{1,t}\}$, where the vertices x_1 and y_1 have no common neighbors in Q . Let $\alpha_1(Q)$ be the graph derived from Q by deleting edges $y_1 y_{1,1}, y_1 y_{1,2}, \dots, y_1 y_{1,t}$ and attaching new edges $x_1 y_{1,1}, x_1 y_{1,2}, \dots, x_1 y_{1,t}$. We say that $\alpha_1(Q) = Q'$ is a α_1 -transform of Q (see Figure 1).

Lemma 1. *Let $\alpha_1(Q) = Q'$ be a tree derived from Q by α_1 -transform as depicted in Figure 1, then*

$$F(\alpha_1(Q)) > F(Q). \quad (6)$$

For any $p > 1, t \geq 1$.

Proof. Observe that $d_{Q'}(x_1) = d_Q(x_1) + t > d_Q(x_1)$ and $d_{Q'}(x_1) + d_{Q'}(y_1) = d_Q(x_1) + d_Q(y_1) = p + t + 2$. Consider that

FIGURE 1: α_1 - transform applied to Q .

$$\begin{aligned}
F(Q') - F(Q) &= \sum_{i=1}^p \left[\left(d_{Q'}(x_{1,i})^2 + d_{Q'}(x_1)^2 \right) - \left(d_Q(x_{1,i})^2 + d_Q(x_1)^2 \right) \right] \\
&\quad + \sum_{j=1}^t \left[\left(d_{Q'}(y_{1,j})^2 + d_{Q'}(x_1)^2 \right) - \left(d_Q(y_{1,j})^2 + d_Q(y_1)^2 \right) \right] \\
&\quad + \left(d_{Q'}(x_1)^2 + d_{Q'}(y_1)^2 \right) - \left(d_Q(x_1)^2 + d_Q(y_1)^2 \right) \\
&= \sum_{i=1}^p \left[\left(d_{Q'}(x_{1,i})^2 + (p+t+1)^2 \right) - \left(d_Q(x_{1,i})^2 + (p+1)^2 \right) \right] \\
&\quad + \sum_{j=1}^t \left[\left(d_{Q'}(y_{1,j})^2 + (p+t+1)^2 \right) - \left(d_Q(y_{1,j})^2 + (t+1)^2 \right) \right] \\
&\quad + \left((p+t+1)^2 + (1)^2 \right) - \left((p+1)^2 + (t+1)^2 \right) \\
&= 3pt(t+p+2) > 0.
\end{aligned} \tag{7}$$

The α_1 - transform decreases the degree of y_1 by t and increases the degree of x_1 by t , while the degrees of the nodes $x_{1,1}, x_{1,2}, \dots, x_{1,p}$ and $y_{1,1}, y_{1,2}, \dots, y_{1,t}$ remain unchanged. \square

Lemma 2. Let $\alpha_2(Q) = Q'$ be a tree derived from Q as depicted in Figure 2, where $d_Q(z_1, u) \geq 1$. Then

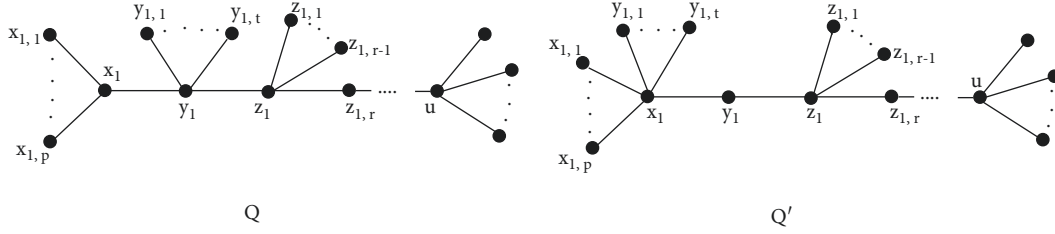
$$F(\alpha_2(Q)) > F(Q). \tag{8}$$

For any $p > 1$ and $t \geq 1$.

Proof. Since $d_Q(x_1) < d_{\alpha_2(Q)}(x_1)$ and $d_{\alpha_2(Q)}(y_1) < d_Q(y_1)$, we have

$$\begin{aligned}
F(Q') - F(Q) &= \sum_{i=1}^p \left[\left(d_{Q'}(x_{1,i})^2 + d_{Q'}(x_1)^2 \right) - \left(d_Q(x_{1,i})^2 + d_Q(x_1)^2 \right) \right] \\
&\quad + \sum_{j=1}^t \left[\left(d_{Q'}(y_{1,j})^2 + d_{Q'}(x_1)^2 \right) - \left(d_Q(y_{1,j})^2 + d_Q(y_1)^2 \right) \right] \\
&\quad + \left(d_{Q'}(x_1)^2 + d_{Q'}(y_1)^2 \right) - \left(d_Q(x_1)^2 + d_Q(y_1)^2 \right) + \left(d_{Q'}(y_1)^2 + d_{Q'}(z_1)^2 \right) - \left(d_Q(y_1)^2 + d_Q(z_1)^2 \right) \\
&= \sum_{i=1}^p \left[1 + (p+t+1)^2 - 1 - (p+1)^2 \right] \\
&\quad + \sum_{j=1}^t \left[1 + (p+t+1)^2 - 1 - (t+2)^2 \right] + \left[(p+t+1)^2 + 2^2 - ((p+1)^2 + (t+2)^2) \right] \\
&\quad + \left[(r+1)^2 + 2^2 - ((t+2)^2 + (r+1)^2) \right] \\
&= 3t[p(t+p+2) - (t+3)] > 0.
\end{aligned} \tag{9}$$

Hence, the result holds. \square

FIGURE 2: α_2 – transform applied to Q .

Lemma 3. Let $\alpha_3(Q) = Q'$ be a tree obtained from Q by applying α_3 -transform (see Figure 3), where $d_Q(z_1, u) = d_{\alpha_3(Q)}(z_1, u) \geq 0$ and $d_Q(x_1, y_1) = d_{\alpha_3(Q)}(x_1, y_1) \geq 2$. If $t \geq 1$ and $s > 1$ then

$$F(\alpha_3(Q)) > F(Q). \quad (10)$$

Proof. By definition of $F(Q)$ we get

$$\begin{aligned}
F(Q') - F(Q) &= \sum_{i=1}^p \left[\left(d_{Q'}(x_{1,i})^2 + d_{Q'}(x_1)^2 \right) - \left(d_Q(x_{1,i})^2 + d_Q(x_1)^2 \right) \right] \\
&\quad + \sum_{j=1}^t \left[\left(d_{Q'}(y_{1,j})^2 + d_{Q'}(x_1)^2 \right) - \left(d_Q(y_{1,j})^2 + d_Q(y_1)^2 \right) \right] \\
&\quad + \left(d_{Q'}(x_1)^2 + d_{Q'}(v)^2 \right) - \left(d_Q(x_1)^2 + d_Q(v)^2 \right) + \left(d_{Q'}(y_1)^2 + d_{Q'}(z_1)^2 \right) - \left(d_Q(y_1)^2 + d_Q(z_1)^2 \right) \\
&\quad + \left(d_{Q'}(y_1)^2 + d_{Q'}(w)^2 \right) - \left(d_Q(y_1)^2 + d_Q(w)^2 \right) = p \left[(p+t+1)^2 + 1 - (p+1)^2 \right] \\
&\quad + t \left[(p+t+1)^2 + 1 - (t+1)^2 \right] + \left[(p+t+1)^2 + 2^2 \right] \\
&\quad - \left[(p+1)^2 + 2^2 \right] + \left[2^2 + 2^2 \right] - \left[2^2 + (t+2)^2 \right] + \left[2^2 + (r+1)^2 \right] - \left[(t+2)^2 - (r+1)^2 \right] \\
&= 3t \left[p(p+t+2) - (t+3) \right] > 0.
\end{aligned} \quad (11)$$

Hence the proof is complete. \square

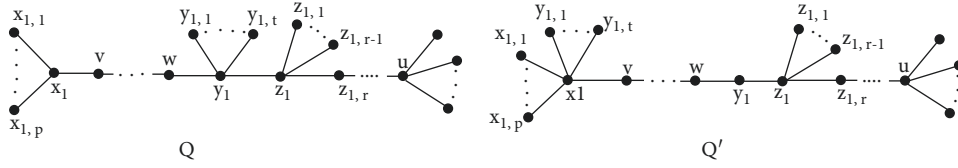
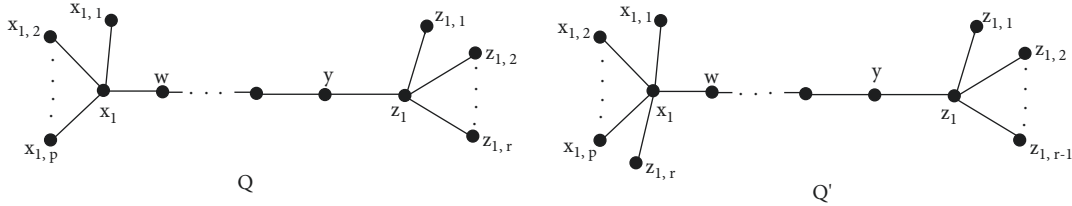
Lemma 4. Let $\alpha_4(Q)$ be a tree obtained from Q after applying α_4 -transform (see Figure 4). For any $p, r \geq 0$, we have

$$F(\alpha_4(Q)) > F(Q). \quad (12)$$

Proof. If $d_Q(x_1, z_1) \geq 2$, then $d_{\alpha_4(Q)}(x_1) + d_{\alpha_4(Q)}(z_1) = d_Q(x_1) + d_Q(z_1) = p + r + 2$. Now by using the definition of F index, we have

$$\begin{aligned}
F(Q') - F(Q) &= \sum_{i=1}^p \left[\left(d_{Q'}(x_{1,i})^2 + d_{Q'}(x_1)^2 \right) - \left(d_Q(x_{1,i})^2 + d_Q(x_1)^2 \right) \right] \\
&\quad + \sum_{j=1}^{r-1} \left[\left(d_{Q'}(z_{1,j})^2 + d_{Q'}(z_1)^2 \right) \right] - \sum_{j=1}^r \left[\left(d_{Q'}(z_{1,j})^2 + d_{Q'}(z_1)^2 \right) \right] \\
&\quad + \left[\left(d_{Q'}(z_{1,r})^2 + d_{Q'}(x_1)^2 \right) \right] + \left[\left(d_{Q'}(x_1)^2 + d_{Q'}(w)^2 \right) \right] - \left[\left(d_Q(x_1)^2 + d_Q(w)^2 \right) \right] \\
&\quad + \left[\left(d_{Q'}(y)^2 + d_{Q'}(z_1)^2 \right) \right] - \left[\left(d_Q(y)^2 + d_Q(z_1)^2 \right) \right] \\
&= p \left[(p+2)^2 + 1^2 - (p+1)^2 - 1 \right] + (r-1) \left[r^2 + 1 \right] - r \left[(r+1)^2 + 1 \right] \\
&\quad + \left[1 + (p+2)^2 \right] + \left[(p+2)^2 + 2^2 \right] - \left[(p+1)^2 + 2^2 \right] + \left[2^2 + r^2 \right] - \left[2^2 + (r+1)^2 \right] \\
&= 3 \left[p(p+3) - r(r+1) + 2 \right] > 0.
\end{aligned} \quad (13)$$

\square


 FIGURE 3: α_3 – transform applied to Q .

 FIGURE 4: α_4 – transform applied to Q .

2.1. Greatest Value of F – Index for Trees of Given Diameter.

The multistar graph denoted by $MS(r_1, r_2, \dots, r_{d-1})$, where $r_1, r_{d-1} \geq 1$ and for $2 \leq j \leq d-2$, $r_j \geq 0$, is the caterpillar involving a path a_1, a_2, \dots, a_{d-1} of length $d-2$ having r_j pendant vertices that are adjacent to a_j for $1 \leq j \leq d-1$. The diameter of $MS(r_1, r_2, \dots, r_{d-1})$ is equal to d , and can be derived by connecting the centers of $K_{1,r_1}, K_{1,r_2}, \dots, K_{1,r_{d-1}}$ with edges. A bistar graph of order r denoted by $BS(p, t)$, where $p + t = r - 2$, is formed by connecting the central vertices of $K_{1,p}$ and $K_{1,t}$ by an edge. A tree that has diameter 3 is also a bistar. For integers r, t with $2 \leq t \leq r-1$, $S_{r,t}$ is tree derived by connecting $t-1$ pendant vertices to the end node of the path P_{r-t+1} , with diameter $d(S_{r,t}) = r - t + 1$.

Theorem 1. *Let T be a tree on $r \geq 3$ vertices and diameter $d \geq 2$. Then the maximum value of $F(T)$ is attained for $T \cong S_{r,r-d+1}$.*

Proof. Applying α_1 -transform on the vertices that are not attached on the diametral path of T , we get that the maximum value of $F(T)$ is attained in the class of multistars $MS(r_1, r_2, \dots, r_{d-1})$. Now applying the transformations presented in Lemma 2–Lemma 4, it follows that the maximum value of $F(T)$ is attained if and only if $r_1 = r - d, r_2 = r_3 = \dots = 0$ and $r_{d-1} = 1$. Hence $T \cong S_{r,r-d+1}$. \square

Corollary 1

(i) *In the set of trees T on r vertices, we have*

$$\max_{d(T)=i} F(T) > \max_{d(T)=j} F(T). \quad (14)$$

(ii) *for $2 \leq i < j \leq r-1$.*

(iii) *In the set of trees T of order r and diameter d with $3 \leq d \leq r-2$, the graphs with the greatest $F(T)$ value are (in this order) as follows:*

$$MS(r-d, 0, \dots, 0, 1), MS(r-d-1, 0, \dots, 0, 2), \dots,$$

$$MS\left(\left\lceil \frac{r-d+1}{2} \right\rceil, 0, \dots, 0, \left\lfloor \frac{r-d+1}{2} \right\rfloor\right). \quad (15)$$

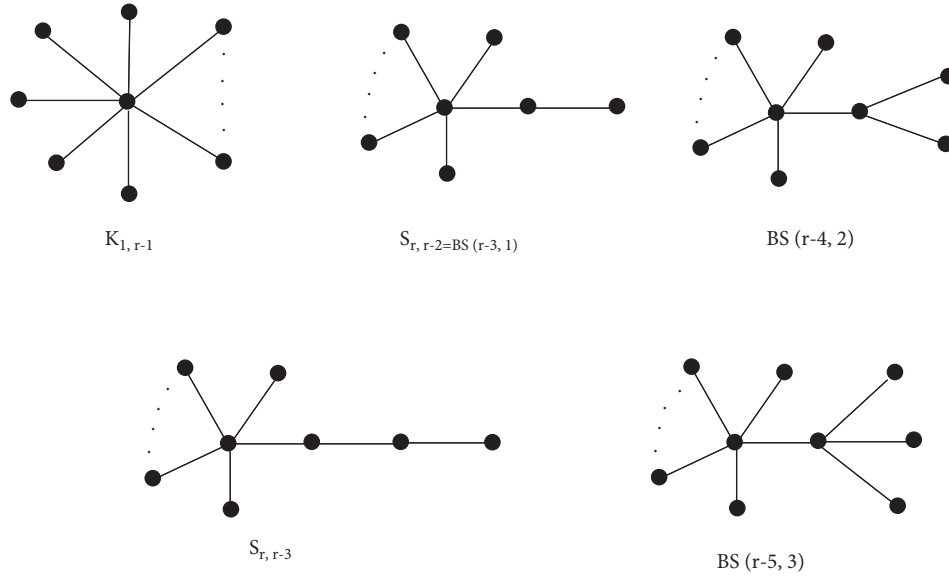
Proof

- (i) Let T be a tree on r vertices with diameter i . By Theorem 1 the maximum value of $F(T)$ is attained for $T \cong S_{r,r-i+1} \cong MS(r-i, 0, \dots, 1)$. The result follows by applying many times α_1 -transform on $MS(r-i, 0, \dots, 1)$.
- (ii) Applying Lemma 1–Lemma 3 to T yields the multistar $MS(p, 0, \dots, 0, q)$ with $p + q = r - d + 1$. Now using Lemma 4 to $MS(p, 0, \dots, 0, q)$, we get the required ordering. \square

Theorem 2. *For tree of order $r \geq 8$, the maximum value of F -index is attained in the following order (see Figure 5).*

$$F(K_{1,r-1}) > F(BS(r-3, 1)) > F(BS(r-4, 2)) > F(S_{r,r-3}) > F(BS(r-5, 3)). \quad (16)$$

Proof. Let T be tree of order $r \geq 8$. By Corollary 1, the maximum value of $F(T)$ is achieved in the set of trees of diameter 2. It follows that the trees with the maximum value of $F(T)$ is star $K_{1,r-1}$. The second maximum is attained for $S_{r,r-2} \cong BS(r-3, 1)$ in the set of trees of diameter 3. The next maximum is reached by $BS(r-4, 2)$ and $BS(r-5, 3)$ in trees of diameter 3. Since $BS(r-3, 1)$ can be obtained from $BS(r-4, 2)$ by a α_4 transformation, we get $F(K_{1,r-1}) > F(BS(r-3, 1)) > F(BS(r-4, 2))$. In the set of trees of diameter 4, the maximum value is attained by $S_{r,r-3}$.

FIGURE 5: Trees T achieving greatest F - index.

To get the fourth maximum value we compare $F(BS(r-5,3))$ with $F(S_{r,r-3})$. We have

$$\begin{aligned} F(BS(r-5,3)) - F(S_{r,r-3}) &= [(r-5)\{1+(r-4)^2\} + \{(r-4)^2+16\} + 3(17)] \\ &\quad - [(r-4)\{1+(r-3)^2\} + \{(r-3)^2+4\} + 8+5] = -3r^2 + 21r + 12 < 0. \end{aligned} \quad (17)$$

Hence $F(BS(r-5,3)) < F(S_{r,r-3})$ for every $r \geq 8$. Also,

$$\begin{aligned} F(BS(r-5,3)) - F(MS(r-5,0,2)) &= [(r-5)\{1+(r-4)^2\} + \{(r-4)^2+16\} + 3(17)] \\ &\quad - [(r-5)\{(r-4)^2+1\} + \{(r-4)^2+4\} + 13+20] = 30 > 0. \end{aligned} \quad (18)$$

This shows that the second maximum value of F -index is achieved by $MS(r-5,0,2)$ after $S_{r,r-3}$ in the set of trees of diameter 4. Using a α_1 -transform, it is easy to see that $MS(r-5,0,0,1)$ reaches the maximum value in the set of

trees of diameter 5 and $F(MS(r-5,0,0,1)) < F(MS(r-5,0,2))$, which completes the proof. \square

Example 1. Let T be a tree on 10 vertices, then

$$\begin{aligned} F(K_{1,10-1}) &= \sum_{uv \in E(Q_{1,10-1})} (d^2(u) + d^2(v)) = 9(1^2 + 9^2) = 738, \\ F(BS(10-3,1)) &= \sum_{uv \in E(BS(10-3,1))} (d^2(u) + d^2(v)) = [1^2 + 2^2] + [2^2 + 8^2] + 7[1^2 + 8^2] = 528, \\ F(BS(10-4,2)) &= \sum_{uv \in E(BS(10-4,2))} (d^2(u) + d^2(v)) = 2[1^2 + 3^2] + [3^2 + 7^2] + 6[1^2 + 7^2] = 378, \\ F(S_{10,10-3}) &= \sum_{uv \in E(S_{10,10-3})} (d^2(u) + d^2(v)) = [1^2 + 2^2] + [2^2 + 2^2] + [2^2 + 7^2] + 6[1^2 + 7^2] = 366, \\ F(BS(10-5,3)) &= \sum_{uv \in E(10-5,3)} (d^2(u) + d^2(v)) = 3[1^2 + 4^2] + [4^2 + 6^2] + 5[1^2 + 6^2] = 288. \end{aligned} \quad (19)$$

This shows that

$$\begin{aligned} F(K_{1,10-1}) &> F(BS(10-3, 1)) > F(BS(10-4, 2)) \\ &> F(S_{10,10-3}) \\ &> F(BS(10-5, 3)). \end{aligned} \quad (20)$$

Theorem 3. *In the set of trees T of order $r \geq 5$ with t pendant vertices, where $3 \leq t \leq r-2$, we have*

$$F(T) \leq t^3 - 7t + 8r - 8. \quad (21)$$

$$F(T) - F(T-x) = (d(y)^2 + 1) - \sum_{z \in N(y) \setminus \{x\}} [d(z)^2 + (d(y)-1)^2 - (d(z)^2 + d(y)^2)]. \quad (22)$$

Since, $d(z_0) \geq 2$ and for the remaining $d(y) - 2$ nodes $z \in N(y) \setminus \{x, z_0\}$, where $d(z) \geq 1$, we have

$$F(T) - F(T-x) = (3d(y)^2 - 3d(y) + 2). \quad (24)$$

We also get $d(y) \leq t$ because $T-y$ includes of $d(y)$ trees. Now $2 \leq d(y) \leq t$ gives

$$F(T) - F(T-x) \leq (3t^2 - 3t + 2), \quad (25)$$

with equality holds if $d(y) = t$, the adjacent vertex of y has degree 2 and the remaining vertices are of degree 1. Hence $T = S_{r,t}$ and x is adjacent to one of the vertex of $S_{r,t}$ of degree t .

Now we use induction to prove the required result. If $r = 5$, then $t = 3$ and we have a bistar $BS(2, 1)$ (see Figure 5), the only tree of order 5 having 3 pendant vertices. Let $r \geq 6$ and suppose the result is true for all trees of order $r-1$ and t pendant vertices, where $3 \leq t \leq r-3$. Let x be a pendant vertex adjacent to y , then we have two cases: (i) y has degree 2 and (ii) y is of degree at most 3.

- (i) The only vertex z adjacent to y has degree $d(z) \geq 2$. Then

$$F(T) - F(T-x) = (d(z)^2 + 4) + 5 - (d(z)^2 + 1) = 8. \quad (26)$$

- (ii) In this case the graph $T-x$ has t pendant vertices. Applying induction, for $t \leq r-3$, we get $F(T-x) \leq F(S_{r-1,t})$, with equality holds for $T-x = S_{r-1,t}$. It follows that

$$F(T) = F(T-x) + 8 \leq F(S_{r-1,t}) + 8 = F(S_{r,t}), \quad (27)$$

- (iii) with equality holds for $T-x = S_{r-1,t}$. Therefore, we have $T = S_{r,t}$. If $t = r-2$, then $T-x$ is a star with one vertex of degree 1. Hence, $T = S_{r,r-2} = S_{r,t}$.

- (iv) Let $T-x$ is of order $r-1$ with s $t-1$ pendant vertices. If $d(y) \geq 3$, then by using induction on $T-x$, we get

The above equality holds if and only if $T = S_{r,t}$.

Proof. First we prove that if x is a pendant vertex adjacent to y , then

$$F(T) - F(T-x) \leq 3t^2 - 3t + 2. \quad (22)$$

With equality holds if and only if $T = S_{r,t}$ and $d(y) = t$. Since $3 \leq t \leq r-2$, it follows that there exists a vertex $z_0 \in N(y) \setminus \{x\}$, such that $d_{z_0} \geq 2$. Otherwise T is a star having central vertex y . We obtain

$$\begin{aligned} F(T) &\leq F(T-x) + 3t^2 - 3t + 2 \\ &\leq F(S_{r-1,t-1}) + 3t^2 - 3t + 2 = F(S_{r,t}), \end{aligned} \quad (28)$$

with equality holds for $T-x = S_{r-1,t-1}$ and $d(y) = t$. Hence $T = S_{r,t}$. \square

3. Conclusion

In this paper, our main focus is to obtain the greatest value of F -index for trees of given order, diameter, and pendent vertices. Also, we determine the ordering of corresponding extremal trees for F -index. We make use of some graph transformations to determine the greatest value of F -index for trees of given order, diameter, and pendent vertices. These transformations involves contraction and swapping of pendant edges from one vertex to other resulting in increase in the value of the forgotten index. Using these transformations continuously on a graph lead us to the desired extremal graph with respect to the forgotten index. On the way of obtaining extremal graphs, we also obtained some other members having the second, third, fourth, and fifth maximum value of the forgotten index.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, "Vertex-and edge-weighted molecular graphs and derived structural descriptors," *Topological Indices and Related Descriptors in QSAR and QSPR*, vol. 169, 1999.

- [2] I. Gutman and N. Trinajstić, "Graph theory and molecular orbitals. Total φ -electron energy of alternant hydrocarbons," *Chemical Physics Letters*, vol. 17, no. 4, pp. 535–538, 1972.
- [3] A. Miličević, S. Nikolić, and N. Trinajstić, "On reformulated Zagreb indices," *Molecular Diversity*, vol. 8, no. 4, pp. 393–399, 2004.
- [4] Y. Yang, H. Liu, H. Wang, and H. Fu, "Subtrees of spiro and polyphenyl hexagonal chains," *Applied Mathematics and Computation*, vol. 268, pp. 547–560, 2015.
- [5] B. Furtula, I. Gutman, and S. Ediz, "On difference of Zagreb indices," *Discrete Applied Mathematics*, vol. 178, pp. 83–88, 2014.
- [6] H. Wiener, "Structural determination of paraffin boiling points," *Journal of the American Chemical Society*, vol. 69, no. 1, pp. 17–20, 1947.
- [7] B. Zhou and N. Trinajstić, "On general sum-connectivity index," *Journal of Mathematical Chemistry*, vol. 47, no. 1, pp. 210–218, 2010.
- [8] B. Furtula and I. Gutman, "A forgotten topological index," *Journal of Mathematical Chemistry*, vol. 53, no. 4, pp. 1184–1190, 2015.
- [9] S. S. Ali, E. R. Kenawy, F. I. Sonbol et al., "Pharmaceutical potential of a novel chitosan derivative Schiff base with special reference to antibacterial, anti-biofilm, antioxidant, anti-inflammatory, hemocompatibility and cytotoxic activities," *Pharmaceutical Research*, vol. 36, no. 1, p. 5, 2018.
- [10] M. Azari and F. Falahati-Nezhed, "Some results on forgotten topological coindex," *Iranian Journal of Mathematical Chemistry*, vol. 10, no. 4, pp. 307–318, 2019.
- [11] Z. Che and Z. Chen, "Lower and upper bounds of the forgotten topological index," *MATCH Commun. Math. Comput. Chem*, vol. 76, no. 3, pp. 635–648, 2016.
- [12] A. Ghalavand and A. R. Ashrafi, "On forgotten coindex of chemical graphs," *MATCH Commun. Math. Comput. Chem*, vol. 83, pp. 221–232, 2020.
- [13] I. Gutman, A. Ghalavand, T. Dehghan-Zadeh, and A. R. Ashrafi, "Graphs with smallest forgotten index," *Iranian Journal of Mathematical Chemistry*, vol. 8, no. 3, pp. 259–273, 2017.
- [14] I. Gutman, K. ChDasChDas, B. Furtula, E. Milovanović, and I. Milovanović, "Generalizations of Szőkefalvi Nagy and Chebyshev inequalities with applications in spectral graph theory," *Applied Mathematics and Computation*, vol. 313, pp. 235–244, 2017.
- [15] I. Tomescu and S. Kanwal, "Ordering trees having small general sum- connectivity index," *MATCH Commun. Math. Comput. Chem*, vol. 69, pp. 535–548, 2013.

Research Article

Development of Water Culture Tourism of Mountain Ethnic Culture Based on 3D Image Technology

Chuangle Guo  and Xinlu Shi

Chengdu University of Information Technology, Chengdu, Sichuan 610000, China

Correspondence should be addressed to Chuangle Guo; chuanglewh@cuit.edu.cn

Received 29 March 2022; Revised 21 April 2022; Accepted 26 April 2022; Published 23 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Chuangle Guo and Xinlu Shi. 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 study of human settlement environment, especially the complete study of human settlement environment in mountainous region, is a huge systematic project, which involves almost all aspects of human knowledge system at the semantic level. It is an important task to establish a systematic cognitive model of human settlement environment and carry out practical verification. In recent years, due to the rise of tourism real estate, mountain tourism real estate, as a type of tourism real estate, has attracted people's attention, and its hidden economic, social, and ecological benefits make it develop rapidly. The purpose of this study is to improve the theoretical system of tourism real estate landscape planning and mountain landscape resource protection system with 3D image technology, which has important theoretical and practical significance for exploring the sustainable development method of mountain tourism real estate landscape. This study locates the research type of mountain tourism real estate, and it is the first time to study mountain tourism, explores the specific planning of mountain tourism real estate development measures, and preliminary establish the theoretical system of mountain tourism real estate development. It brings a new idea for the development of mountain tourism real estate and puts forward the direction and method of planning.

1. Introduction

With the development of the economy and the improvement of people's living standards, many buyers are simultaneously turning to the tourism industry, which has rapidly spread from large- and medium-sized cities to small-sized cities. Therefore, vacation has become the main form of tourism in a short time [1]. In order to meet the demands of consumers, real estate has seized the opportunity. A large number of resorts and hotels have been established in scenic resorts along the coast [2]. However, too fast development tends to lead to problems such as imperfect infrastructure, poor conceptual overall planning, and destructive landscape ecological environment. An unsustainable economic growth model and consumption model will inevitably bring severe challenges and threats to

human ecological environment and social development. Especially for the tourism industry, the impact is very far-reaching [3, 4].

Mountainous area is a terrain with uneven surface distribution and elevation difference reaching a certain threshold. However, from the perspective of human social attributes, mountain characteristics are actually an important cause of spatial differentiation [5]. [6, 7]. . Of course, this reason is not direct. No matter from the perspective of a spatial scale, the spatial differentiation of social people is generally manifested as concentration or dispersion on a large spatial scale, thus deriving the urban living environment. People create living environment, and living environment affects people's behavior. This feedback is particularly strong in mountain man environments [6,7]. From the present single perspective, the ultimate problem of

human settlement is to further study the relationship between spatial heterogeneity and human settlement construction, so as to realize the harmonious coexistence and sustainable development of mountain environment and human system. The settlement environment here is not a city alone, but an organic community composed of regional cities, towns, communities, and even buildings [8, 9].

With the continuous development of tourism real estate, how to carry out the construction of landscape environment of tourism real estate, to achieve the harmonious balance and sustainable development of economy and social culture by landscape ecology, is a problem that we are facing at present and need to solve. Tourism real estate development is a collection of landscape environment and the combination of tourism real estate planning cross domain [10]. For now, in regard to the development of tourism real estate in our country, we mainly consider three aspects. One aspect is the economic interest. The developer is the main purpose of the development in order to obtain the maximum economic return. This is cannot change. Another aspect is focused on tourism real estate planning study, mainly producing industry planning. The third aspect is the landscape planning research of tourism real estate itself, mainly studying the landscape ecological planning of tourism real estate. However, considering these aspects alone, we cannot carry out effective and reasonable development and construction [11]. We must stand on the basis of the overall landscape environmental benefits and take ecology, landscape science, environmental landscape ecological sustainable development of tourism geography culture, and other related theories, an overall evaluation of the tourism real estate development [12]. The purpose of this study is to discuss how to carry out the tourism real estate landscape development approach according to the environmental conditions of specific regions through relevant theories [13].

New theories promote the development and evolution of new technologies, and new technologies often, in turn, promote the development and innovation of theories. The main technology platform of this study is based on 3D image processing technology, which is a beneficial attempt to innovate the discipline methodology of architectural planning. This study tries to find a new scientific research model of human settlement environment and supports the scientific and expansibility of human settlement environment with new methods and solid scientific logic. As far as the current situation is concerned, 3D image technology has been applied in the field of urban planning, but this kind of research lacks an integrated, systematic, and theoretical research perspective, especially the research on the cultural tourism of famous ethnic groups in mountain areas, which makes it possible for this research to achieve some original results. The research significance of this study is summarized as follows:

- (1) With the development of tourism real estate, the infrastructure of tourism resort is constantly built and improved, and a large number of construction are poured into coastal, lake, mountain, and other

landscape tourism areas, resulting in the destruction of many original landscape environment. At the same time, due to the over-development of tourism real estate in the local area, the local traditional social culture gradually declined or even disappeared. Therefore, to maintain continuation and development of landscape culture resources, we must first consider the future development of tourism real estate is the problem; only through reasonable and sustainable development of protection principles can we make our eternal vitality and unique tourism real estate landscape.

- (2) In the 1990 s, China began to tourism real estate in 1992. The State Council approved the construction of 12 state-level tourist resorts, and so far, various types of tourist real estate have developed rapidly across the country. However, compared with developed countries such as Europe and the United States , tourism real estate in China started late; the related theoretical system is relatively backward; in order to establish perfect theory system, we must combine landscape science, environmental, ecology, and other related theory, to explore suitable for China's national conditions of tourism real estate development real principles and methods, so as to promote the perfection of the theory on tourism real estate development in China [16].
- (3) The development of tourism real estate will certainly promote the development of local economy, which is of course beyond doubt. However, we also want to scientifically use of reasonable planning and improve the utilization rate of resources, to maximize the local resource advantage to tourism as today's popular industry, not only brought the rapid development of economy but also benefits the local people, and a more solid foundation has been laid for the construction of socialist avenue [17, 18].

2. Related Work

Tourist real estate in the south of France and Naples, Italy, had its earliest forms before the nineteenth century. In 1839, Lord Brougham, the British Prime Minister, built his own villa on the Bay in southern France and praised it, thus arousing the great interest of most British people in the bay. In 1858, the French poet, Conneau, sent a letter to the wife of Napoleon III and once again made the southern coast of France a holiday hot. Due to the establishment of villas and hotels and various entertainment facilities, the most famous leisure and health resorts are made here in Europe, thus attracting many nobles to gather here. After World War I (1918), before the advent of the Great Depression, the traditional Puritan morality collapsed and the hedonism of the jazz age became popular. The value of this land on the south coast of France was once again favored by Americans, and American jazz culture replaced the original aristocratic coastal vacation culture [19].

In the early 1980s, the concept of timeshare vacation exchange began to spread all over the world, mainly developed capitalist countries in Europe. During this period, a large number of investors and developers in this area, due to the increasing product, make product business model more clear, relevant law also constantly improves and forms a complete set of services, and management industry and exchange industry also developed rapidly mature; according to relevant data, in 1999, the global timeshare property sales was as high as 6.72 billion dollars, and the United States accounted for 55%, among which the sale of rooms accounted for 50.4%. The world's largest resort, Condominium International (RCI), is an American company founded in 1974. Headquartered in Indiana, USA, the company is the first in the world to introduce the concept of timeshare exchange. As one of the global vacation rental market leader, RCI has more than 3 million members worldwide and covers 73,000 resorts worldwide [20, 21]. In the past three years, RCI vacation membership has also increased by 25 percent. In recent years, Japan, South Korea, India, Thailand, Philippines, Singapore, Malaysia, and other Asian countries have also played a tourism holiday brand. Resort hotels, leisure, health, leisure, hot spring clubs, weekends, holidays, etc., are also the heart of the development of the world economy. International financial investors and hotel management and investment institutions have also been involved, making the forward development of tourism real estate more hopeful [22].

China's tourism real estate started late, starting with timesharing resort hotels in Hainan in the 1990s. With the development of China's reform and opening up, Hainan has its unique tropical never winter island, beautiful scenery, pleasant climate, and fresh air, which attracts a large number of real estate developers and presents the real estate demand in Hainan. During this period, Hainan's tourism real estate accounted for 40% of the national market share. Hainan's real estate development mode integrating leisure, and vacation tourism is overturning the traditional real estate development mode. Hainan has always been regarded as a paradise for tourism and vacation. Tourism real estate started earlier [23]. However, after the real estate bubble in 1992 and 1993, tourism real estate was really taken seriously vacant commercial housing [20]. Starting in 1999, Hainan has spent seven years of processing after the bubble burst Commercial housing backlog. In the whole process, tourism real estate has played a big role This road is proved that the tourism real estate development model not only can make the Hainan real estate out of the shadows of the bubble but also make it step-by-step prosperity. In 2005, Hainan closely combined real estate and tourism with tourism real estate as the main brand [24], creating a peak of tourism real estate development in Hainan. Hainan has always been regarded as a paradise for tourism and vacation, and tourism real estate started earlier. However, after the real estate bubble in 1992 and 1993, tourism real estate was really taken seriously. The bubble economy has been the pain of Hainan real estate for a long time, which directly brought about a large number of half-completed projects and idle commercial houses. In the

whole process, tourism real estate has played a significant role. This road proves that the development mode of tourism real estate can not only make Hainan real estate out of the shadow of bubble but also make it gradually to prosper. In 2005, Hainan closely combined real estate and tourism, with tourism real estate as the main brand, creating a peak of tourism real estate development in Hainan [25, 26].

Figure 1), [27]. . According to the characteristics of the surface, the surface can be divided into four basic forms: plain, hill, plateau, and mountain (as shown in Figure 1)1. The surface has a large undulating area, mostly in the area of tectonic movement and external force activity, and the geological structure is complex, which is characterized by large absolute height and relative height [27]. Mountain has broad sense and narrow sense, general mountain, highland mountain, and hill three parts, among which narrow sense mountain only refers to very low mountain. Since the measurement process of spatial uniformity itself is a continuous cycle, the research object of this study is the generalized mountain tourism [28].

On the contrary, while the theory and practice of human settlement environment science are flourishing, as a cross-sectional methodology, the related theory and empirical research of 3D image technology is also flourishing in the geoscience, although the research of 3D image mountain tourism is also in a stage of exploration and research. It can be noted that, so far, the relevant research on mountain tourism is mainly focused on the natural field (landform, water system, natural landscape, etc.). From the perspective of semantic model, the object and category of mountain cultural tourism research are quite different from that of human settlement environment research, which can only be used for reference in methodology [29]. Therefore, the research of human settlements info-tupu is inherently exploratory and has certain risks. However, it is precisely because of the interdisciplinary nature of the academic field that it is easiest to achieve fruitful results between scientific fields. It is worth noting that there are still some areas for improvement in the current research:

- (1) Parallel: the first is environment science and the information mapping theory of parallel; so far, the living environment of scientific research achievements of theory and practice has widely appeared, geological information of map-related research also gradually prospers [30].
- (2) Object discretization: at present, the outstanding problem of human settlements environment-related research is object discretization, and there is lack of effective integration. The research methods are flat and descriptive, and there is no effective methodology and technical model for the study of human environment system. The object discretization is mainly manifested by the fuzzy definition of the object of human environment system or the semantic repetition and logical crossover in the definition of research objects.

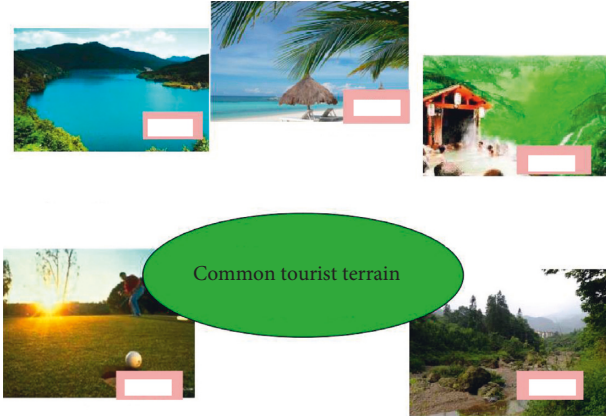


FIGURE 1: The common tourist terrain.

- (3) Mountain plains: the main problem is the lack of summary of the characteristics and rules of living environment; just the plain human settlement environment model and the design method to be used to study mountain problems [31]. The grid layout of mountain cities has become popular in recent years. Therefore, this study proposes the research on water culture of mountain ethnic culture tourism development based on 3D image technology [31].

The contribution of this study is as follows:

- (1) Improve the theoretical system of tourism real estate landscape planning and mountain landscape resource protection system with 3D image technology
- (2) This work has important theoretical and practical significance for exploring the sustainable development method of mountain tourism real estate landscape

3. 3D Image Technology-Based Water Culture of Mountain Ethnic Culture Tourism

3.1. Water Culture of Mountain Ethnic Culture Tourism. The research is a typical cross-cutting research, which has the following key issues to be solved at the theoretical and technical level: (1) semantic (conceptual) composition and logical relationship sorting (object) of (mountain) human settlement environment system; (2) the constitution and internal contradiction of human settlement system (structure); (3) identification and expression model of human settlement environment system (mountain) (technology); (4) index system and numerical modeling (method) of (mountain) human settlement environment; (5) (mountain) regional verification of spatial information Atlas of mountain human settlement environment (test); (6) law condensing and model innovation (summary). To analyze and express the living environment through the spectrum of spatial information graph of the living environment, its key logical order is as follows: clear the research object, analyze the object structure, establish the expression model, establish the analysis model, establish the empirical model, and propose the planning model.

Among the six key links, the most important problem is the living environment of semantic model because when the former human settlement environment still stays on the concept description, it is an open system; combined with the related discipline, each holds one word and is not easy to use unified agreed to the key words to describe key logic structure of living environment. As a result, the solutions to these key problems are largely subjective. In the practical process, it is necessary to combine theory with practice, in order to carry out comprehensive research and analysis on the basis of the existing theoretical and practical achievements. At the technical level, the study takes 3D image technology as the platform, comprehensively integrates various research methods, and forms a research framework of mountain culture tourism.

3.2. 3D Image Technology. Convolutional neural network is one of the important branches of deep learning. The super-resolution technology based on convolutional neural network mainly relies on the rich operations of neural network such as convolution, deconvolution, pooling, and activation. Convolutional layer is the basis of neural network. Through the convolutional kernel (convolutional template), input signals can be extracted from different levels, depths, and even different frequency bands of information. The convolutional kernel is similar to various operator templates in traditional algorithms, such as gradient operator and smoothing operator. However, traditional template operator values are fixed, and specific features can only be extracted according to specific settings. Different from this, convolutional templates in neural networks only need to be initialized and then automatically updated through the iterative optimization of the network. At present, the convolutional kernels are divided into 2D convolution and 3D convolution, and their sizes are generally odd.

In the network design architecture, in addition to the linear or nonlinear processing of data, there are many specialized normalized layer operations that are helpful for network training convergence. For example, the layer normalizes data in channel direction, and the group norm divides channels. The most effective method is Batch Norm (BN), which refers to subtracting the mean and dividing by the variance in the ordinary sense, while BN is to normalize the data of eachbatch. In addition to subtracting the mean and dividing the variance, scale transformation parameter γ and offset parameter are also introduced. This can not only make the output data of the upper layer conform to normal distribution but also keep the characteristic structure of the output of the upper layer unchanged. The following is the derivation of the specific formula:

$$u = \frac{1}{N} \sum_{i=1}^N x_i, \quad (1)$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - u)^2, \quad (2)$$

$$\hat{x}_i = \frac{x_i - u}{\sqrt{\sigma^2 + \epsilon}}, \quad (3)$$

$$y_i = \gamma \hat{x}_i + \beta. \quad (4)$$

The above super-resolution method is to build some basic layers through the network, based on these basic operational layers. At present, convolutional neural networks, used to achieve image super-resolution, can be roughly divided into series and parallel networks, residual network, pyramid network, recursive network, dense connection network, and generative adversarial network and the combination between them. The structure of series and parallel networks is relatively simple and the input of low resolution rate graph. For example, the serial convolutional structure of the network should not be too deep through the direct output of stacked convolutional layers. It is an efficient structure and has significant advantages over other structures, whether used for classification tasks such as recognition or regression tasks such as super-resolution reconstruction. The selection of network training optimizer mainly includes batch gradient descent (BGD), stochastic gradient descent (SGD), minibatch gradient descent (MBGD), and momentum optimization algorithm. The momentum learning method weighs the importance of the previous gradient and the current gradient on the basis of SGD, and the formula is expressed as

$$W_t = W_{t-1} + \alpha V_t, \quad (5)$$

$$V_t = \beta V_{t-1} + (1 - \beta) dW. \quad (6)$$

The RMSprop optimizer can adaptively adjust the learning rate by using the gradient accumulation of previous training, and its updating formula is

$$W_t = W_{t-1} + \frac{\alpha}{\sqrt{\theta_i + \epsilon}} dW, \quad (7)$$

$$\theta_i = \beta \theta_{i-1} + (1 - \beta) (dW)^2. \quad (8)$$

Adam optimizer combines a compromise version of momentum optimization and RMSprop optimization, which is the most stable among all optimizers at present. Its optimization strategy can be summarized as

$$g = \frac{1}{m} \sum_{i=1}^m \frac{\partial L(y^i, f(x^i, w))}{\partial w}, \quad (9)$$

$$v_t = \beta_1 v_{t-1} + (1 - \beta_1) g, \quad (10)$$

$$r = \beta_2 r + (1 - \beta_2) g^2, \quad (11)$$

$$W_t = W_{t-1} + \frac{\alpha}{\sqrt{r + \epsilon}} v_t. \quad (12)$$

Based on equations (1)–(12), Figure 2 gives 3D image technology-based water culture of mountain ethnic culture tourism proposed in this study.

4. Experimental Results and Analysis

4.1. Introduction to Experimental Dataset. This study mainly adopts the method of corpus research and combines quantitative statistics with qualitative analysis. First of all, we collected news texts from People's Daily for consecutive months and established a large database. Then, according to the research purposes of different chapters, we extracted samples from the main corpus and built a secondary corpus. For some chapters, we also classified and collected them according to the needs. Then, we sorted all the subcorpora in each chapter of the newspaper into text documents to establish a pair of subcorpora.

4.2. Experimental Results' Analysis. There are many kinds of tourism resources, including viewing and recreation resources, ecological environment resources tourism services, facilities and service level, hotel catering, leisure entertainment, etc. All kinds of tourism resources are leisure tourism areas developed by vacationers or tourists and are the material basics for entertainment, sports, fitness, and other activities. The type, quality, and scale of tourism resources directly determine the attraction and development potential of the resort. Through the transformation of buildings, vegetation, and water bodies in Shinan Water culture tourism zone, the efficiency of water-bearing forest is improved. The dust-holding capacity of rivers is conducive to the degradation and dilution of concentration of some pollution sources and the improvement of water ecosystem function and water environment quality. The comparison before and after the transformation is shown in Figure 3.

The size of the derived agents (artificial and social) of the corresponding human settlements also increases. In a long period of history, the suitability state of mountain tourism area is approximately unchanged, that is, during this period, the adaptability of human settlement environment is mainly passive adaptation. If the object of human settlement is simplified as topographic and geomorphic environment, the suitability of the object of human settlement can be approximately regarded as a fixed value. With the development of human society, especially after the arrival of industrialization process (usually also manifested as rapid urbanization process), the characteristics of the subject and object of human settlement environment have undergone tremendous changes. The focus on hotspot distribution before (right) and after (left) mountain water resource rehabilitation design project is given in Figure 4.

From Figure 5, we know the topographic and geomorphic characteristics of the planning area (the attribute domain is uncertain). The planning area is located in the deep hill and low mountain area, with broken topography, large terrain slope, lack of large flat land, poor urban land conditions, and low land utilization rate. The coefficient of buildable land in the planning land is about 40–50%, and the elevation distribution characteristics is of the planning area (attribute field is 1251). The altitude span of the planning area is large, ranging from 115 to 1366 meters at the lowest level. Combined with the topography and geomorphology, it

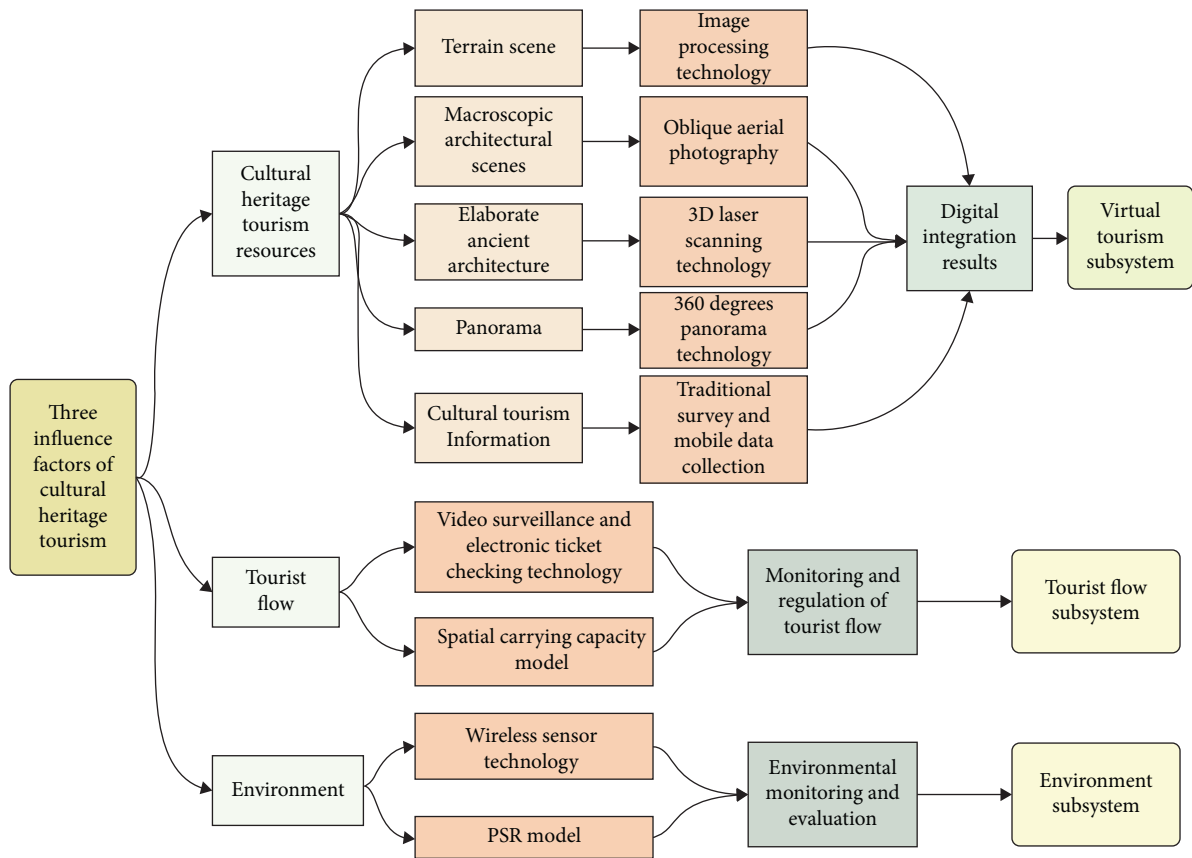


FIGURE 2: 3D image technology-based water culture of mountain ethnic culture tourism framework.

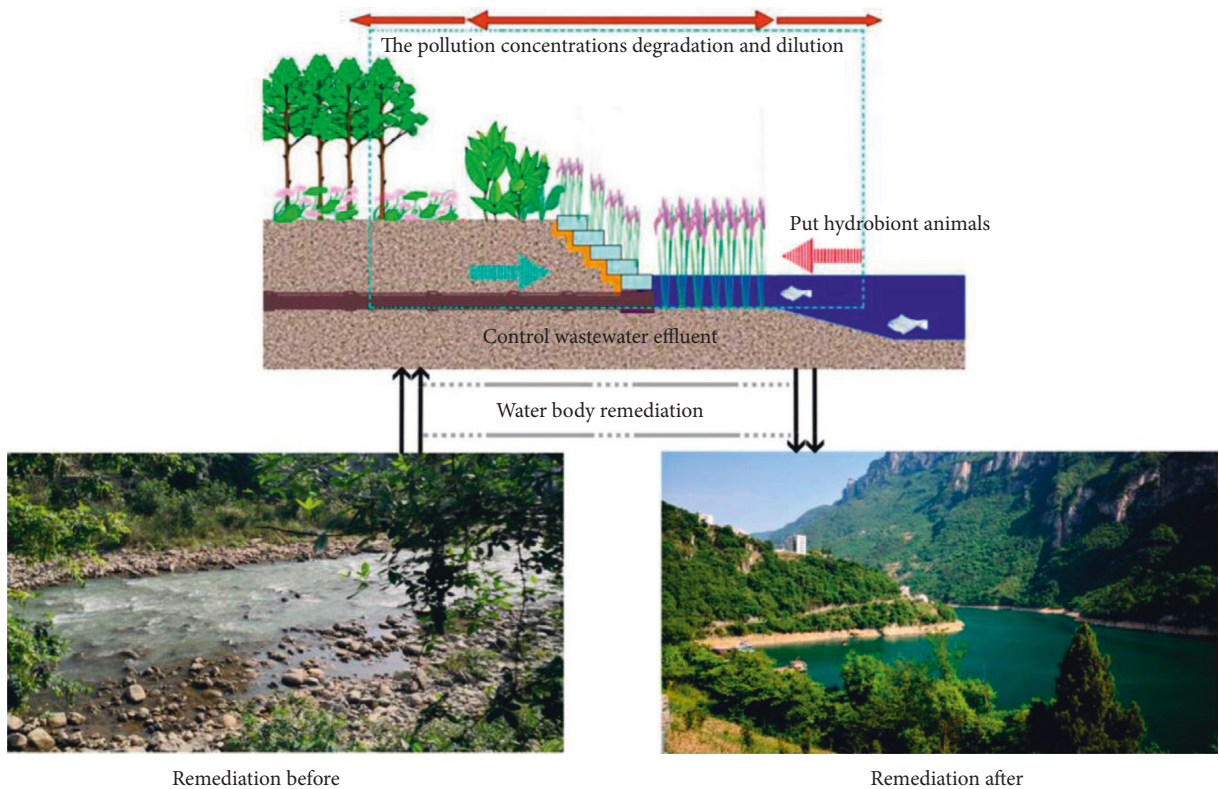


FIGURE 3: Visualization results of water restoration in mountain tourism area.

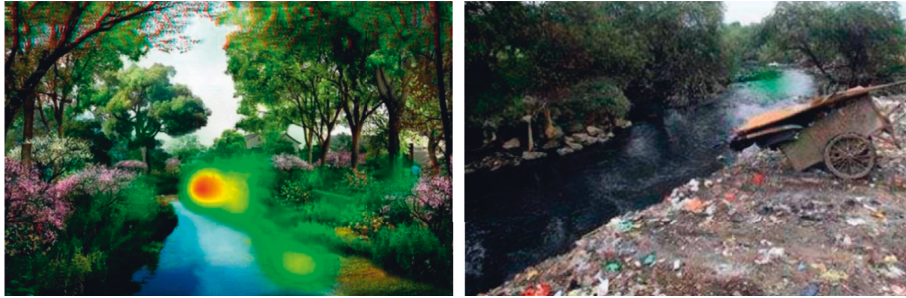


FIGURE 4: Focus on hotspot distribution before (b) and after (a) mountain water resource rehabilitation design project.

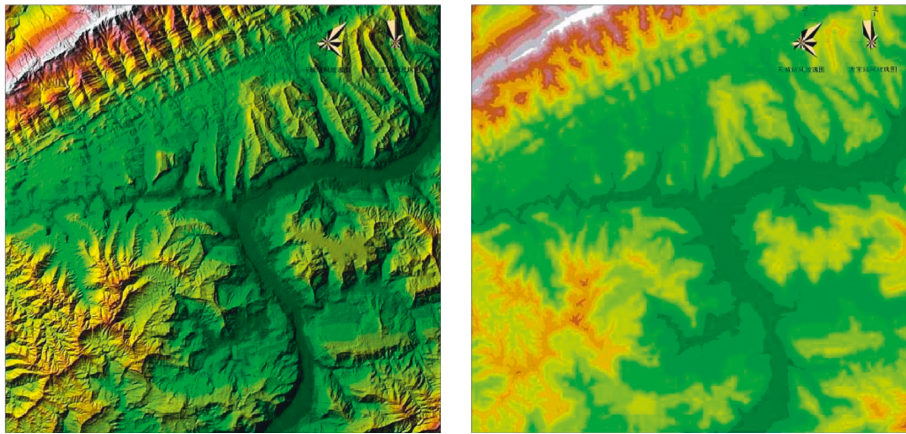


FIGURE 5: Landform space form (a) and Spatial form of elevation distribution (b).

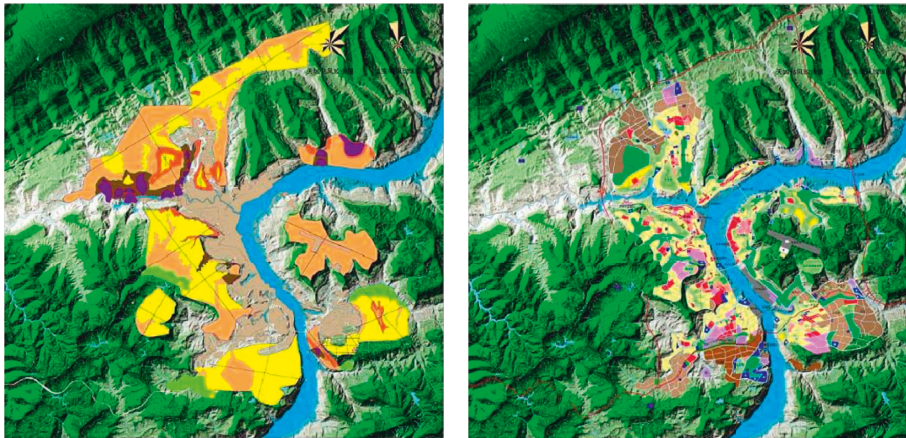


FIGURE 6: Space form of engineering geological conditions (a) and spatial form of land use planning (b).

presents an obvious platform elevation (M) ranging from 115–215, 215–415, 415–615, 615–915, and 915. The proportion above was 11%, 39.1%, 33.4%, 13.4%, and 13%, respectively.

In addition, from Figure 6, we know the characteristics of engineering geological conditions in the planning area (attribute field is 3). There are 18 dangerous rocks and 15 landslides in the planning area which have passed the risk assessment of geological disasters. The planning space is divided into three types: forbidden area, cautious area, and suitable area. Forbidden area is mainly distributed in the upper flat on the north bank of Zhuxi River and the area at

the foot of Lion Mountain, and careful area mainly distributed in the north of ShenMingBa, including the building of the area, is not prone to geological disasters and low rock. The structure of urban land is a multicenter group layout structure, consisting of 8 groups and 3 areas (Longbao Tiancheng Five Bridges). The total planned land area is 80.99 km², and residential land accounts for 25.1% in terms of attribute composition. Public facilities, industrial storage, external transportation, roads, squares, municipal public facilities, green land, and special land accounted for 14.7%, 21.4%, 2.8%, 9.8%, 13.3%, 5.6%, 13.8%, and 2.3%, respectively.

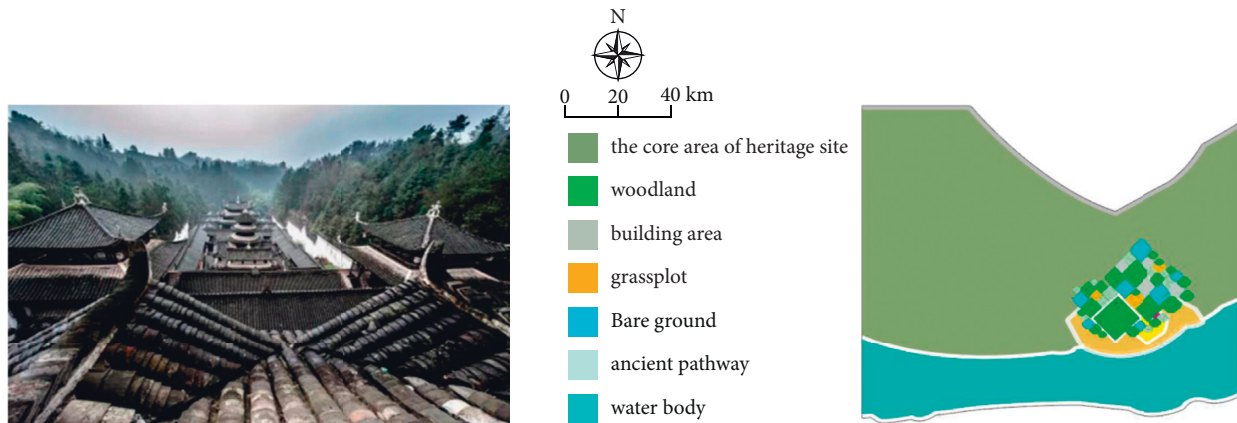


FIGURE 7: Cultural landscape pattern design project after renovation.



FIGURE 8: Exhibition of water culture tourist destination before and after reconstruction.

Due to the unique tourism resources and huge development potential of tourist resorts, it is of great significance to study and explore the development of famous cultural tourism in mountain areas for a hundred years. On the basis of not damaging local landscape resources, this study designed a conceptual effect map, as shown in Figure 7.

Mountain landscape attraction is mainly manifested in the following three aspects, Firstly, the mountain landscape environment is different from Binhai lakeside hot spring tourism real estate of another type, such as relying on the unique mountain natural ecological environment to attract tourists. Secondly, mountain villa holidays, especially in scorching summer, became the most suitable resort mountain. Thirdly, culture and entertainment are the soul of mountain tourism real estate. In order to make the development of mountain tourism real estate obtain lasting attraction, there must be a certain cultural environment, such as folk customs, temples, and historical sites. Entertainment mainly refers to the supporting facilities of tourism sites, such as commercial streets, hotels, and ski mountain bike races, which make tourism real estate sufficiently attractive and serve tourists better and drive the development of local economy. Exhibition of water culture tourist destination before (b) and after (a) reconstruction is given in Figure 8.

5. Conclusions

Through consulting a large number of books and literature, in the comprehensive use of landscape ecology, tourism,

geography, biology, and other multidisciplinary theoretical knowledge, based on the research of 3D image technology, the study explores the related technologies and theories of combining mountain water cultural tourism real estate development with landscape planning and design and makes relevant explorations and studies on the problems existing in related concepts, content, technology, environmental conditions, and other aspects of planning. The research of this study will expand the research scope of mountain culture tourism and has great theoretical value and potential application prospect.

Data Availability

The dataset can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Z. Cai, C. Fang, Q. Zhang, and C Fulong, "Joint development of cultural heritage protection and tourism: the case of Mount Lushan cultural landscape heritage site," *Heritage Science*, vol. 9, no. 1, pp. 1–16, 2021.
- [2] Z. Lin, L. Zhang, S. Tang, Y. Song, and X. Ye, "Evaluating cultural landscape remediation design based on VR technology," *ISPRS International Journal of Geo-Information*, vol. 10, no. 6, p. 423, 2021.

- [3] J. Su and J. Sun, "Spatial changes of ethnic communities during tourism development: a case study of Basha Miao minority community," *Journal of Tourism and Cultural Change*, vol. 18, no. 3, pp. 333–350, 2020.
- [4] Y. Deng, S.-Y. Han, J. Li, J. Rong, W. Fan, and T. Sun, "The design of tourism product CAD three-dimensional modeling system using VR technology," *PLoS One*, vol. 15, no. 12, Article ID e0244205, 2020.
- [5] Y. Sun, M. Wang, Y. Zhang, and Z. Chun, "Application of the innovative design and digital technology in the construction of traditional village digital museum," in *Proceedings of the 2020 international conference on innovation design and digital technology (ICIDDT)*, pp. 354–357, Zhenjiang, China, 2020.
- [6] I. A. Chandio, A. N. B. Matori, K. B. Wanyusuf, M. A. H. Talpur, S. H. Khahro, and M. R. M. Mokhtar, "Computer application in routing of road using least-cost path analysis in hillside development," *Research Journal of Environmental and Earth Sciences*, vol. 4, no. 10, pp. 907–911, 2012.
- [7] P. Tarolli and E. Straffelini, "Agriculture in hilly and mountainous landscapes: threats, monitoring and sustainable management," *Geography and Sustainability*, vol. 1, no. 1, pp. 70–76, 2020.
- [8] H. Gao, W. Xu, X. Wang, and Z. Qiang, "Research on digital protection of bamboo tube based on hani nationality," *IEEE, in Proceedings of the 2021 7th international conference on computer and communications (ICCC)*, pp. 960–964, Chengdu, China, December 2021.
- [9] X. Deng, I. T. Kim, and C. Shen, "Research on convolutional neural network-based virtual reality platform framework for the intangible cultural heritage conservation of China hainan Li nationality: boat-shaped house as an example," *Mathematical Problems in Engineering*, vol. 11423, 2021.
- [10] Y. Zhang, M. Han, and W. Chen, "The strategy of digital scenic area planning from the perspective of intangible cultural heritage protection," *Eurasip Journal on Image and Video Processing*, vol. 2018, no. 1, pp. 1–11, 2018.
- [11] J. Zhang, J. Wang, S. D. Min, K. K. Chen, and H. Huang, "Influence of curriculum quality and educational service quality on student experiences: A case study in sport management programs," *Journal of Hospitality, Leisure, Sport and Tourism Education*, vol. 18, pp. 81–91, 2016.
- [12] X. Guo, W. Jiang, Q. Zhang, and W. Kai, "Digital protection technology of cultural heritage based on ArcGIS geographic information technology algorithm," *Security and Communication Networks*, vol. 2022, pp. 17–26, Article ID 3844626, 2022.
- [13] L. Wu, H. Jiang, W. Chen, and W. Peng, "Geodiversity, geotourism, geoconservation, and sustainable development in xiangxi UNESCO global geopark—a case study in ethnic minority areas," *Geoheritage*, vol. 13, no. 4, pp. 1–18, 2021.
- [14] K. Sang, G. L. Fontana, and S. E. Piovani, "Assessing railway landscape by ahp process with gis: a study of the yunnan-vietnam railway," *Remote Sensing*, vol. 14, no. 3, p. 603, 2022.
- [15] Q. Liu, "Development and design of folk arts intangible cultural heritage products with laser scanning," *Microscope*, vol. 11, pp. 1234–1247, 2021.
- [16] A. I. Ramaano, "The implied significance of integrative geographical information systems in sustainable tourism and comprehensive community development in Musina Municipality, South Africa," *Technological Sustainability*, vol. 26, no. 01, pp. 90–101, 2021.
- [17] D. Hutárová, I. Kozelová, and J. Špulerová, "Tourism development options in marginal and less-favored regions: a case study of Slovakia's gemer region," *Land*, vol. 10, no. 3, p. 229, 2021.
- [18] R. Lama and D. Sarkhel, "Ethnic tourism[M]//Encyclopedia of tourism management and marketing," *Edward Elgar Publishing*, vol. 9, no. 4, pp. 76–89, 2022.
- [19] S. N. R. Wijesinghe, P. Mura, and R. Tavakoli, "A postcolonial feminist analysis of official tourism representations of Sri Lanka on Instagram," *Tourism Management Perspectives*, vol. 36, Article ID 100756, 2020.
- [20] W. D. Sulisty, M. N. Lukmanul Khakim, N. Jauhari, and R. D. Anggraeni, "Fun learning history: explore the history of water sites based on android," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 16, no. 07, pp. 105–118, 2021.
- [21] V. Argyropoulos and A. Stratigea, "Sustainable management of underwater cultural heritage: the route from discovery to engagement-open issues in the mediterranean," *Heritage*, vol. 2, no. 2, pp. 1588–1613, 2019.
- [22] J. Fu, J. Zhou, and Y. Deng, "Heritage values of ancient vernacular residences in traditional villages in Western Hunan, China: spatial patterns and influencing factors," *Building and Environment*, vol. 188, Article ID 107473, 2021.
- [23] T. Jiang, X. Gan, Z. Liang, and L. Guang, "AIDM: artificial intelligent for digital museum autonomous system with mixed reality and software-driven data collection and analysis," *Automated Software Engineering*, vol. 29, no. 1, pp. 1–22, 2022.
- [24] Y. Yang and H. Kang, "The digital measures for protection and heritage of dongba culture," in *Recent Developments in Intelligent Computing, Communication and Devices*, vol. 6, pp. 1203–1208, Springer, Singapore, 2019.
- [25] S.-N. Zhang, Y.-Q. Li, C.-H. Liu, and W.-Q. Ruan, "Does live performance play a critical role in building destination brand equity - a mixed-method study of "Impression Dahongpao"," *Journal of Retailing and Consumer Services*, vol. 59, Article ID 102392, 2021.
- [26] R. Gagne, J. B. Barreau, F. Lécuyer, T. Nicolas, N. Jean-Marie, and V. Gouranton, "eXtended reality for cultural heritage," pp. 67–72, 2022.
- [27] T. W. Whyke, J. L. Mugica, and M. S. Brown, "Contemporizing the national style in Chinese animation: the case of nezha (2019)," *Animation*, vol. 16, no. 3, pp. 157–174, 2021.
- [28] J. A. Alzubi, R. Jain, A. Kathuria, A. Khandelwal, A. Saxena, and A. Singh, "Paraphrase identification using collaborative adversarial networks," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 1, pp. 1021–1032, 2020.
- [29] T. A. Nguyen, B. T. Nguyen, H. Van Ta et al., "Livelihood vulnerability to climate change in the mountains of Northern Vietnam: comparing the Hmong and the Dzaio ethnic minority populations," *Environment, Development and Sustainability*, vol. 23, no. 9, Article ID 13469, 2021.
- [30] P. Mansilla-Quiñones, H. Manríquez, and A. Moreira-Muñoz, "Virtual heritage: a model of participatory knowledge construction toward biogeocultural heritage conservation," *Global Geographical Heritage, Geoparks and Geotourism*, pp. 75–94, Springer, Singapore, 2021.
- [31] D. T. Huyen, D. T. N. Huy, and N. T. S. Linh, "Potential for the development of community tourism in vo nhai site (Thai nguyen - vietnam)," *Revista Gestão Inovação e Tecnologias*, vol. 11, no. 3, pp. 1355–1370, 2021.

Research Article

Performance Evaluation Model of Agricultural Enterprise Technology Innovation Based on GA-BP Neural Network

Jian Kang^{1,2} and Minjuan Zhao ¹

¹College of Economics and Management, Northwest A&F University, Yangling, Shaanxi 712100, China

²Center for Shaannan Eco-Economy Research, Ankang University, Ankang, Shaanxi 725000, China

Correspondence should be addressed to Minjuan Zhao; minjuan.zhao@nwsuaf.edu.cn

Received 26 March 2022; Revised 19 April 2022; Accepted 31 May 2022; Published 23 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Jian Kang and Minjuan Zhao. 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 my country's economy and society, how to effectively evaluate the technical innovation performance of agricultural enterprises has become the focus of research. This paper firstly processes and analyzes the technical innovation performance data of agricultural enterprises, and then processes and converts the technical innovation performance data of agricultural enterprises; then, through the analysis of the technical innovation performance data of agricultural enterprises, the key characteristics of the technical innovation performance of agricultural enterprises are excavated. Finally, a performance evaluation model based on agricultural enterprise technology innovation is proposed, and the validity of the model is verified with examples.

1. Introduction

At present, more than 60% of agricultural invention patents and more than 80% of new agricultural product development in China are completed by agricultural enterprises. With the continuous deepening of social and economic development led by scientific and technological innovation in my country, how to evaluate the innovation performance and management level of different types of scientific and technological innovation entities has become a major management problem to be solved urgently in scientific research and innovation. Develop a scientific, fair, and practical performance appraisal system and build an information-based and data-based management platform, which is helpful for the functional orientation of the main body of agricultural scientific research and innovation, problem-oriented mining, and development decision-making. With the continuous deepening of social and economic development led by scientific and technological innovation in my country, how to evaluate the innovation performance and management level of different types of scientific and technological innovation entities has become

a major management problem to be solved urgently in scientific research and innovation. Develop a scientific, fair, and practical performance appraisal system and build an information-based and data-based management platform, which is helpful for the functional orientation of agricultural scientific research and innovation subjects, problem-oriented mining, and development decision-making. Small and medium-sized enterprises are not only effective carriers to accelerate the transformation of scientific and technological achievements and realize technological innovation but also an important source of national economic growth. Therefore, how to objectively, scientifically, and effectively evaluate the performance evaluation of agricultural enterprises' technological innovation is of great significance for promoting the development of agricultural enterprises' technological innovation and improving their core competitiveness. The concept of "technological innovation" was first put forward by economist Schumpeter, who defined innovation as establishing a new production function [1]. Since the 1980s, many scholars at home and abroad have conducted extensive research on technological innovation. In 1985, Musser made a systematic and overall analysis of

technological innovation [2]. Technological innovation is defined as technological innovation based on the novelty of its conception and its successful realization of meaningful discontinuous events as features [3]. Fu Jiayi believes that the technological innovation capability of an enterprise refers to the conditions and power of technological development and transformation that an enterprise exhibits in the process of technological innovation activities [4]. At present, there are many methods for evaluating the technological innovation capability of enterprises, among which the widely used methods include the Delphi method, AHP, principal component analysis, grey system evaluation, data envelopment analysis, and fuzzy comprehensive evaluation. [5]. However, their disadvantage is that the random factors in the evaluation have more influence on the evaluation results, and the evaluation results are easily affected by the subjective consciousness, experience, and knowledge limitations of the evaluators [6, 7]. Back-propagation neural network is currently the most widely researched and applied artificial neural network in various fields. It embodies the most essential part of artificial neural network theory and application [8]. The technological innovation capability of an enterprise is affected by many factors. These influences are not isolated. They are interrelated and restrict each other, forming a complex nonlinear system [9]. As an effective tool to solve nonlinear system problems, this paper establishes a GA-BP neural network evaluation model to comprehensively evaluate the technological innovation ability of small and medium-sized enterprises [10].

The goal of technological innovation performance evaluation is to use scientific evaluation methods to measure and evaluate the completion of the established goals of agricultural enterprises based on selected comprehensive indicators and to objectively evaluate an enterprise through the results and put forward constructive suggestions [11, 12]. The main evaluation methods are the balanced scorecard method, the analytic hierarchy process, and other linear regression analysis methods, which are quite mature. Different from general enterprise performance evaluation, in the evaluation of main financial indicators, technological innovation performance evaluation has certain particularity. Du [13] first pointed out that priority should be given to improving inventory turnover rate and accounts receivable turnover rate. The weights of indicators that fit with the characteristics of agriculture are adapted to the characteristics of agriculture. Literature [14] established a performance evaluation system of agricultural enterprise technology innovation about lean manufacturing with the help of the balanced scorecard method and proposed an indicator system of four dimensions: customer, shareholder, production, and continuity. Wu [8] constructed a complete set of indicators for agricultural performance evaluation using grounded theory and proposed a measurement system for agricultural performance. The human brain is a very complex network structure, that is to say, the human brain completes intelligence, emotions, and other advanced psychological activities through a complex network structure, which is the original source of neural networks [15, 16].

Neural network is a technology that abstracts and simulates the general features of the human brain or nature [17–19]. In recent years, neural networks have achieved remarkable results in the field of prediction and empirical research. Compared with traditional prediction models, BP neural network has unique advantages [20] and has been widely used in enterprise performance, financial risk, and bank risk assessment. Many scholars have carried out effective combination innovation of BP neural network and enterprise evaluation system and verified the validity of the evaluation system by using the neural network model. Zuo et al. [21] used the neural network model to evaluate the breakthrough innovation performance of enterprises and analyzed and concluded that the BP neural network has good generalization ability in performance evaluation. Also using the BP neural network, Li et al. [22] constructed and verified the evaluation system of my country's listed banks. Li et al. [23] conducted research on private enterprises in my country, established a performance evaluation index system including enterprise objectives, partnership, and internal process, and combined neural network and dynamic fuzzy method to demonstrate the evaluation system. Other scholars such as Du [13, 14] applied the BP neural network model to evaluate the application.

The technical innovation performance evaluation of agricultural enterprises is affected by many factors [9]. These influences are not isolated. They are interrelated and restrict each other to form a complex nonlinear system. Therefore, based on the BP neural network model and combined with the genetic algorithm, this paper proposes the genetic algorithm-backpropagation neural network to comprehensively evaluate the technological innovation ability of small and medium-sized enterprises [24, 25]. First, analyze and process the technical innovation performance information of agricultural enterprises, and then construct a performance-based evaluation index system for small and medium-sized enterprises' technological innovation capabilities. Considering the entire process of production and operation, it includes two categories of indicators: resource input capability and innovation capability. Finally, combining BP neural network with the genetic algorithm, a performance evaluation model of agricultural enterprise technological innovation based on the GA-BP neural network is proposed. The identification method of electricity stealing behavior of energy Internet users based on GA-BP neural network can effectively strengthen the evaluation of technological innovation performance of agricultural enterprises. The research is aimed at different types of agricultural enterprises, on the basis of determining the positioning of agricultural enterprises' technological innovation performance evaluation, focusing on scientific and technological innovation, technology promotion, platform construction, talent team building and management level, select targeted and representative indicators. Build an information-based and data-based evaluation management platform to achieve rational evaluation and scientific decision-making.

2. Data Analysis and Processing of Technical Innovation Performance Evaluation of Agricultural Enterprises

Data mining refers to the science that extracts hidden laws from a large amount of data through systematic analysis and predicts the future or guides future work according to these laws. It is a process of extracting hidden and potentially useful information and knowledge [26, 27]. Data mining includes problem mining and raising, data processing, data mining analysis, and result evaluation. The process is shown in Figure 1.

- (1) *Problem Mining and Posing*. Through the analysis of the problems encountered in practical work and the corresponding data, the problem to be solved is proposed, and the appropriate mining algorithm is selected [28, 29]
- (2) *Data Processing*. Preprocessing the data according to the characteristics of the algorithm
- (3) *Data Mining Analysis*. Data mining algorithms are used to process data and dig out potentially valuable information

2.1. Data Processing of Technical Innovation Performance Evaluation of Agricultural Enterprises. In the technical innovation performance evaluation data, there are data missing caused by the acquisition. If these values are deleted, it will seriously affect the technical innovation performance evaluation and analysis results. Therefore, the missing part of the technical innovation performance evaluation data is filled with the Lagrangian interpolation method to ensure the technical innovation data enterprises. After realizing the frame-by-frame operation, the signal is windowed. In order to clearly show the effect, a rectangular window $w(n)$ is selected. The calculation formula is

$$L_n(x) = \sum_{i=0}^n l_i(x)y_i, \quad (1)$$

$$l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j},$$

where x is the sequence number in the table below. $L_n(x)$ is the added interpolation result, and x_i is y_i the subscript sequence number of the original value.

2.2. Selection of Performance Evaluation Indicators for Technological Innovation of Agricultural Enterprises and System Construction. The "Performance Excellence Evaluation Criteria" was released in August 2004 by referring to the evaluation criteria of the American Markom Porich National Quality Award and combining the actual situation of my country's quality management. Starting from seven major modules, including the standard comprehensively, comprehensively, and systematically analyzes the entire

process of enterprise operation, and all the factors involved to become an organization. Pursue standards of quality management and excellent performance. As a systematic and comprehensive evaluation criterion, the standard of excellent performance evaluates the quality management level of the enterprise and the overall performance of the enterprise. The overall level of an enterprise as a system includes not only the level of enterprise quality management and enterprise performance but also the operation level and innovation ability of the enterprise. As a reflection of the overall system level of the enterprise, the technological innovation capability of the enterprise can also be evaluated according to the performance excellence standard, and the entire process of enterprise operation can be analyzed and researched more comprehensively. The performance evaluation indicators of listed household appliance enterprises are listed in Table 1.

3. Identification Method of Electricity Stealing Users in Energy Internet Based on GA-BP Neural Network

A genetic algorithm (GA) is a global search evolutionary algorithm that refers to the biological evolution process. Analogous to the biological evolution process, the genetic algorithm simulates biological evolution through selection, inferiority, and mutation to generate next-generation solutions. The heuristic optimal solution search method of the genetic algorithm combined with the BP neural network can effectively solve the problem that the BP neural network is easy to fall into the local optimal solution, and improve the accuracy of the technical innovation performance evaluation of agricultural enterprises. The specific GA-BP network process is shown in Figure 2.

The GA-BP neural network process is as follows.

3.1. Genetic Algorithm Initialization. Before using the genetic algorithm, it is necessary to decode the solution data in the solution space into genotype string structure data. The commonly used encoding method is the binary encoding method and different string structure data form different points. The N string-structured data generated by the encoding represent different individuals. Through the representation of different individuals, the birth population is finally formed, also known as the initial population.

3.2. Fitness Evaluation. Fitness indicates the individual's ability to adapt and also indicates the individual's pros and cons. The calculation formula of the individual's adaptability is

$$F = \text{mse}(Y - O) = \frac{1}{n} \sum_{i=1}^n (y_i - o_i)^2, \quad (2)$$

$$F = \text{mse}(X - O) = \frac{1}{m} \sum_{i=1}^m (x_i - |o_i|)^2,$$

where m is the total number of samples; x_i is the output result of the genetic algorithm; o_i is the actual output result



FIGURE 1: Data mining process.

of the genetic algorithm, and mse represents the mean square error function.

3.3. Genetic Manipulation. The principle is applied to the population selection of genetic algorithm, and excellent individuals are selected for genetic reproduction of the next generation. Individuals with strong adaptability can pass good genes to offspring through inheritance. In this paper, the proportional selection strategy is adopted, the population data size is set as M , and then, the fitness formula of individual i is

$$p_i = \frac{k/F_i}{\sum_{i=1}^N k/F_i}, \quad (3)$$

where k is the genetic coefficient.

In the genetic algorithm, the next generation of individuals is reproduced through the crossover operation, so as to obtain new individuals with new characteristics. The crossover operation is used to exchange genetic information and the mutant individual, which has the characteristics of small value.

3.4. Combination of Genetic Algorithm and BP Neural Network. Using the optimal value search of the genetic algorithm can reduce the calculation amount of the BP neural network and optimize the calculation process of the BP neural network. The fault tolerance of the entire network will be deteriorated, resulting in a reduction in the ability to recognize unlearned samples. For the irregular content in the sample, its generalization ability will also be reduced. Previous research theories have confirmed that for a typical 3-layer BP neural network, when the number of input layer units is n and the number of hidden layer units is $2n + 1$, the network can approximate any differentiable with arbitrary precision function. The BP neural network maps the input variables to the output variables nonlinearly through the excitation function and continuously adjusts the weight threshold of each layer connection. The BP neural network draws on the model structure of the neural network and builds a highly complex learning system through the interconnection of a large number of neurons. The hidden layer performs data calculation by setting one or more layers of neurons, and each layer of neurons can have several nodes. Double-hidden BP neural network is commonly used in a pattern classification problem, which has the advantage of fast classification speed, so the BP neural network structure with double hidden layer is adopted in this paper. The neural network is adjusted through the weight matrix and error feedback between the layers to achieve the expected output result of the load. Compared with the

traditional artificial neural network, the double hidden layer BP neural network has improved in parallel processing of massive data and accuracy. BP network topology is shown in Figure 3.

The processed data enter the second step of processing, and the frame-by-frame operation is realized by using the method of field overlap. In order to accurately analyze the data, the data are divided into t frames, and the output of the first hidden layer node i is

$$b_i = f \left(\sum_{m=1}^m w_{mi} a_m - \theta_i \right). \quad (4)$$

The output of the second hidden layer node j is

$$c_j = f \left(\sum_{i=1}^I w_{ij} b_i - \theta_j \right). \quad (5)$$

The output result of the n th node of the output layer is

$$d_n = f \left(\sum_{j=1}^J w_{jn} c_j - \theta_n \right). \quad (6)$$

In order to improve the convergence speed of the BP neural network, the input data are normalized to reduce the range of changes and improve the flexibility of interval selection. The formula is as follows:

$$t'_i = \frac{t_i - t_{\min}}{t_{\max} - t_{\min}}, \quad (7)$$

where t_i is the input value of the model, t'_i is the normalized value of the set $[0, 1]$ interval, t_{\min} and t_{\max} are the upper and lower limits of the set value, respectively.

The activation function is as follows:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (8)$$

In view of the above analysis, through the collection and analysis of user electricity consumption information, a method for identifying electricity stealing behavior of energy Internet users has a very positive effect on the long-term safety.

The analysis process is described as follows:

- (1) Create a database. Analyze, collect, and process all agricultural enterprise technological innovation performance evaluation data during the period.
- (2) Analyze the GA-BP neural network in the database. Collect mining potential information of technical innovation performance evaluation of agricultural enterprises.

TABLE 1: Performance evaluation indicators of listed home appliance manufacturers.

First-level indicator	Secondary indicators	Three-level indicator
Ability to invest resources	Human resources	Number of scientific and technical personnel
		Structure of scientific and technical personnel
	Financial resources	The proportion of scientific and technological personnel in the total number of people
		Total investment in technological innovation
		Technological innovation investment structure
		The proportion of technological innovation investment in sales revenue
Innovation process capability	Information resource	Information resource base construction
	Infrastructure	Number of experiments and test equipment
		Experiment and test equipment structure
	Related party relationship	Investment in the construction of supplier and strategic partnership
		Supplier and strategic partnership maintenance investment
	Innovation value creation process	Identification of the innovation value creation process
Determination of requirements for innovation value creation process		
Innovation output capability	Innovation support process	Design of innovative value creation process
		Implementation and improvement of innovation value creation process
	Quality	Identification of innovation support processes
		Determination of innovation support process requirements
		Design of the innovation support process
		Implementation and improvement of the innovation support process
Innovation environment support	Type	The advanced nature of technical performance indicators of innovation achievements
		Customer satisfaction and loyalty for innovation outcomes
	Difference	Innovation outcome cost structure
		The difference between the cost of innovation and other similar products or substitutes degree
	Finance	The average development cycle of innovation achievements
		Average time between production of innovations
The structure of innovation achievements		
Performance results of innovation outcomes		
Innovation environment support	External environment	Market share of innovative achievements
		Operating income from innovation achievements
	Internal environment	Total profit of innovation achievements
		Contribution rate of total assets of innovation achievements
		Contribution rate of government policies to enterprise innovation
		Total financial investment

- (3) Develop a plan, i.e., plan for the performance evaluation of agricultural enterprises' technological innovation that has been excavated.
- (4) Compare information, i.e., compare the information in the forecast with the known information, and then deploy according to the previously formulated plan.

4. Experimental Analysis

At the beginning of the model iteration, the fitness value of the individual population is far from the optimal fitness, and the individual fitness value increases significantly. In the later stage of the model iteration, due to the continuous convergence of the model, the fitness value of the individual population is

getting more and more close to optimal fitness. After completing the above work, the global optimal initial weight threshold value searched by the genetic algorithm can be obtained, and the initial weight threshold value is brought into the network to train the BP network. Set the training parameters of the BP network: the number of neurons in the input layer is 6, the number of neurons in the hidden layer is 9, and the number of neurons in the output layer is 1. Select the unipolar Sigmoid function as the transmission. For the selection of function, the gradient descent BP training algorithm function training is selected as the training function, where the parameter is set as follows: the maximum number of training times is 200 times, the learning rate is 0.1, and the learning minimum mean square error target is 0.003. According to the selected 12 input vectors, the output data are normalized.

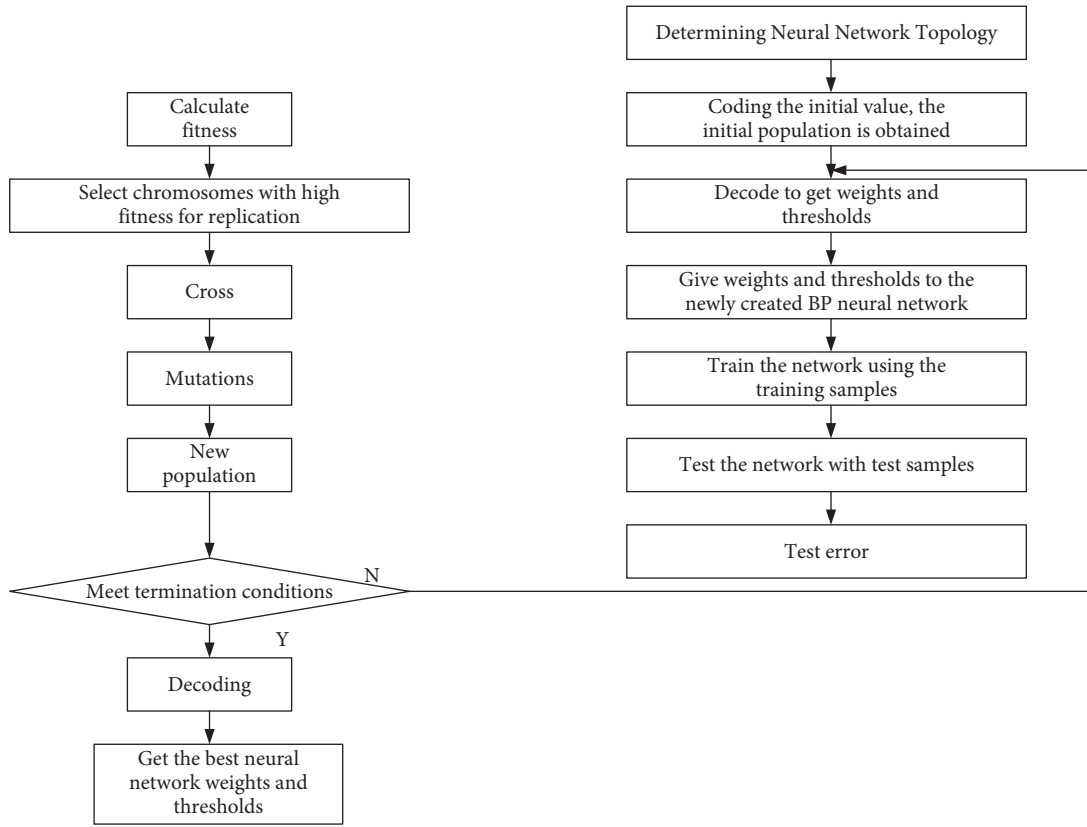


FIGURE 2: GA-BP network flow chart.

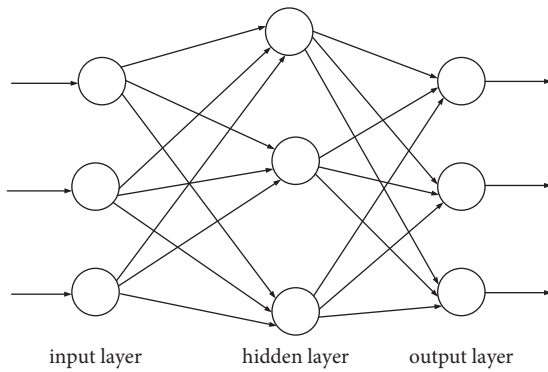


FIGURE 3: BP network topology.

According to the constructed model, a BP neural network model with two hidden layers is established by using Matlab, and the normalized data are input to iterate until the output result is full and the prediction reaches the accuracy required by the experiment. Compared with a BP neural network, GA-BP neural network is closer to the technical innovation performance evaluation of agricultural enterprises.

The results of the experiment can be seen in Figure 4. The error between the predicted value and the actual value of the test sample is shown in Table 2.

By testing the test set composed of the technical innovation performance evaluation data of agricultural enterprises in 2020, the relative error between the test results and

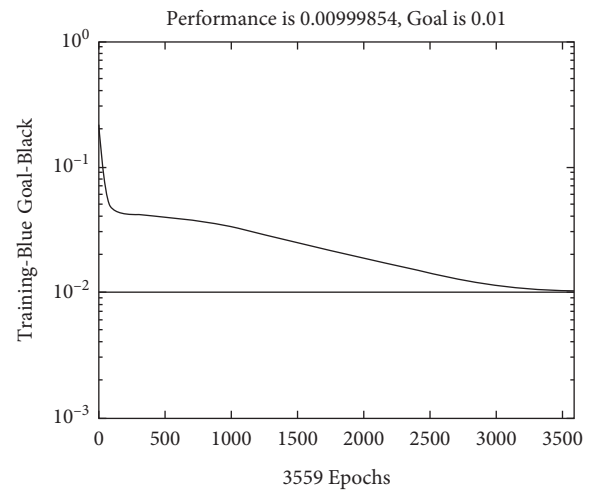


FIGURE 4: Experimental results.

TABLE 2: The error between the predicted value and the actual value.

Serial number	26	27	28	29	30
Deviation	-0.084	0.009	-0.052	0.060	-0.006

the expected results meets the requirements of the target error, which can effectively indicate that the constructed network model has very good performance. It can be seen

that the evaluation results are very close to the actual results, which verifies the feasibility of the GA-BP neural network model in the performance evaluation of agricultural enterprises' technological innovation.

5. Conclusion

This paper proposes a model based on GA-BP neural network which provides a scientific method. The adopted GA-BP neural network algorithm can realize arbitrary linear and nonlinear function mapping, avoid artificial determination of weights, and reduce the evaluation of the evaluators in their cognition. Through systematic training and learning, the system error can meet the accuracy requirements. Research on the performance of agricultural enterprises can solve the problem of fairness evaluation between different types and different responsibilities. Analyzing and evaluating the differences in different aspects of the technological innovation performance of agricultural enterprises from different levels, to find out the gaps between the evaluation subjects, is conducive to development decision-making and development of targeted development plans. The whole set of evaluation index system proposed by the research is suitable for other nonprofit agricultural enterprises and has certain promotion and application value. For nonpublic scientific research institutes and technology enterprises, it is also possible to increase the relevant indicators such as industrial income or adjust the weight of the evaluation indicators on the basis of the evaluation system. The constructed performance evaluation database management system can realize the networkization of performance evaluation workflow management, evaluation index management, and reporting data management, and realize the whole process automation of performance evaluation, calculation and score, data statistics, and result release, and has system expansion functions. These high-quality functions are conducive to the promotion and application of the system. The algorithm overcomes the shortcomings of BP, improves the calculation speed and processing performance compared with the statistical algorithm and dynamic memory algorithm, is more suitable for the task of technological innovation performance evaluation, and provides a new idea for the field of technological innovation performance evaluation of agricultural enterprises.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The authors thank special fund for National Modern Agricultural Industrial Technology System (oats and buckwheat) (no. CARS-07-F-1).

References

- [1] T. Lian, M. Xie, and J. Xu, L. Chen and H. Gao, "BP neural network model is used for odd-even discrimination of integer number," in *Proceedings of the 2013 International Conference on Optoelectronics and Microelectronics (ICOM)*, pp. 67–70, IEEE, Piscataway, New Jersey, United States, September 2013.
- [2] H. Yin and D. You, "Research on the performance evaluation of enterprise breakthrough innovation based on BP neural network," *Soft Science*, vol. 28, no. 05, pp. 125–129+144, 2014.
- [3] Y. Shu and G. Xu, "Multi-level dynamic fuzzy evaluation and BP neural network method for performance evaluation of Chinese private enterprises," *Wireless Personal Communications*, vol. 102, no. 4, pp. 2715–2726, 2018.
- [4] K. Hou, G. Shao, H. Wang et al., "Research on practical power system stability analysis algorithm based on modified SVM," *Protection and Control of Modern Power Systems*, vol. 3, no. 1, p. 11, 2018.
- [5] S. Caihua, L. Yuanjuanm, F. Ting, H. Zhai, and B. Li, "Esearch on customer segmentation model of power nterprise," *Modern Electronics Technique*, vol. 37, no. 2, pp. 91–94, 2014.
- [6] A. de Jong, K. de Ruyter, D. Isobel Keeling, A. Polyakova, and T. Ringberg, "Key trends in business-to-business services marketing strategies: developing a practice-based research agenda," *Industrial Marketing Management*, vol. 93, pp. 1–9, 2021.
- [7] B. L. Hu, J. R. Xu, H. H. Gao, J. H. Liu, and K. R. Wang, "Modified BP neural network model is used for odd-even discrimination of integer number," *Applied Mechanics and Materials*, vol. 423–426, pp. 2675–2678, 2013.
- [8] Y. Wu, X. Li, Q. Liu, and G. Tong, "The analysis of credit risks in agricultural supply chain finance assessment model based on genetic algorithm and backpropagation neural network," *Computational Economics*, pp. 1–24, 2021.
- [9] Q. Wang, X. Wang, Y. Zhao, and H. Tian, "Performance Evaluation of High-Tech Enterprises' Science and Technological Input Based on BP Neural Network," in *Innovative Computing*, pp. 1239–1246, Springer, Berlin/Heidelberg, Germany, 2020.
- [10] X. Wu, J. Liu, and Y. Peng, "A novel heuristic approach for sustainable social and economic development based on green computing technology and big data," *Journal of Enterprise Information Management*, vol. 2021, pp. 1741–0398, 2021.
- [11] G. Chen, J. Shan, D Li et al., "Research on wind power prediction method based on convolutional neural network and genetic algorithm," in *Proceedings of the 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*, pp. 3573–3578, IEEE, Asia, May 2019.
- [12] L. Wen and X. Yuan, "Forecasting CO2 emissions in Chinas commercial department, through BP neural network based on random forest and PSO," *The Science of the Total Environment*, vol. 718, p. 137194, 2020.
- [13] Y. Du, "A Novel Agricultural Engineering Trend Prediction Based on Neural Network Prediction Model with the Rough Set Theory," 2017.
- [14] C. Shi, M. Han, and Y. Ju, "Evaluation of smart city developmental level based on principal component analysis and GA-BP neural network," in *Proceedings of the International Conference on Frontier Computing*, pp. 359–371, Springer, Tokyo, Japan, July 2022.
- [15] P. Sujatha and D. Lalitha Bhaskari, "Sentence wise Telugu to English translation of vemana sathakam using LSTM," *International Journal of Recent Technology and Engineering*, vol. 8, no. 4, pp. 10739–10743, 2019.

- [16] Y. Ruan and N. Jan, "Design of intelligent recognition English translation model based on deep learning [J]," *Journal of Mathematics*, vol. 2022, no. 107, pp. 1–10, 2022.
- [17] Y. Chen and A. Syed Hassan, "Business English translation model based on BP neural network optimized by genetic algorithm," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–10, 2021.
- [18] X. Duan, "The innovative development path of College English translation teaching under the network environment," in *Proceedings of the 2019 3rd International Conference on Advancement of the Theory and Practices in Education (ICATPE 2019)*, pp. 1553–1559, Francis Academic Press, Helsinki, Finland, November 2019.
- [19] H. Jiang, "English translation of traditional Chinese medicine classics from the perspective of German functionalist translation theory," in *Proceedings of the 8th International Conference on Social Science, Education and Humanities Research (SSEHR 2019)*, pp. 87–90, Francis Academic Press, Sanya, China, November 2019.
- [20] C. Chen, "On intercultural education in college English translation teaching," *Frontiers in Educational Research*, vol. 2, no. 7, pp. 61–65, 2019.
- [21] Z. Zuo, S. Li, J. Yang, Y. Yuan, and Li Xiang, "Research on water resources carrying capacity of shale gas development area based on GA-BP neural network," *Journal of Environmental Engineering and Technology*, vol. 11, no. 01, pp. 194–201, 201.
- [22] Z. Li, "The construction of the turning classroom of business English translation teaching in higher vocational education under the Internet + environment," *Frontiers in Educational Research*, vol. 2, no. 5, pp. 36–42, 2019.
- [23] B. Li, K. Zheng, Z. Wang, and L. Jiale, "Service provider portfolio selection for project management using a BP neural network," *Annals of Operations Research*, vol. 308, no. 1, pp. 41–62, 2022.
- [24] H. Zhou, "Application of GA-BP neural network model in enterprise performance evaluation," *Accounting communications*, no. 17, pp. 31–36, 2019.
- [25] N. Y. Yen and J. C. Hung, "Evaluation of smart city developmental level based on principal component analysis and GA-BP neural network," *Frontier Computing*, vol. 422, pp. 359–371, 2018.
- [26] X. Yang, "Credit evaluation of power supply enterprises based on BP neural network optimized by Genetic Algorithm," *New Economy*, no. 11, pp. 61–66, 2021.
- [27] L. Wang and X. Bi, "Risk assessment of knowledge fusion in an innovation ecosystem based on a GA-BP neural network," *Cognitive Systems Research*, vol. 66, pp. 201–210, 2021.
- [28] M. Gen, L. Lin, Y. Yun, and H. Inoue, "Recent advances in hybrid priority-based genetic algorithms for logistics and SCM network design," *omputers & Industrial Engineering*, vol. 125, pp. 394–412, 2018.
- [29] J. Feng, B. Yuan, L. Xin, D. Tian, and W. Mu, "Evaluation on risks of sustainable supply chain based on optimized BP neural networks in fresh grape industry," *Computers and Electronics in Agriculture*, vol. 183, Article ID 105988, 2021.

Research Article

Robust Parameter Design for Cyclone System Based on Dual-Response Surface Method and Multiobjective Genetic Algorithm

Fusheng Luo ^{1,2} Xianhui Yin ³ and Zhanwen Niu¹

¹College of Management and Economics, Tianjin University, Tianjin 300072, China

²School of Management, Hainan University, Haikou, Hainan 570228, China

³College of Quality and Standardization, Qindao University, Qindao, Shandong 266071, China

Correspondence should be addressed to Xianhui Yin; qsyinxianhui@qdu.edu.cn

Received 19 April 2022; Revised 28 May 2022; Accepted 30 May 2022; Published 18 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Fusheng Luo 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 realize the predictive control of coal preparation quality and ensure that the quality of washing products is close to the minimum coal quality requirement of coal blending to the greatest extent is one of the important means to maximize production and maintain the interests of customers and enterprises. Therefore, the feasibility of introducing the double response surface method and multiobjective genetic algorithm to solve the aforementioned problems is further discussed. By selecting the controllable factors and noise factors affecting the output and determining their respective value levels, the product table method is used to design the robust parameter design test, and the experimental results are obtained, according to the experimental data, the second-order polynomial model of the mean and standard deviation of each response characteristic is established, and the effectiveness of the model is analyzed. Then, the double-response optimization function of each response characteristic is established according to the type of response characteristic. Finally, the corresponding parameter values are solved by multi-objective genetic algorithm. The internal and external surface method is used to design and run 60 tests. Through optimization analysis, the robust parameter settings are 150.68 kpa, 0.18143.73 kpa, and 30%, and the optimal output is ash 8.499%, which yields 69.54%, meeting the requirements of stakeholders. Moreover, compared with the traditional optimization design method, the superiority of the proposed method is verified, which shows that this method is conducive to the transformation of the coal preparation plant from fire-fighting quality management to preventive quality management and provides support for the accurate control and systematic management of the production process of the coal preparation plant.

1. Introduction

Coal preparation is the main technological process for the value-added of coal products. It is the key link for coal enterprises to identify customer needs, meet customer needs, and create enterprise value. Therefore, the qualification and stability of product quality and the washing rate in the coal preparation process directly affect the coal preparation effect. If the quality of washing products fluctuates greatly or is out of control, it will not only bring technical problems to the subsequent coal blending and affect the interests of customers but also cause enterprises to pursue too high quality and reduce the washing rate to meet

the requirements of coal blending, so as to bring benefit losses to enterprises. In order to ensure the stability of coal preparation product quality and maximize the washing rate, experts and scholars at home and abroad have conducted a lot of research. As the coal preparation process is the fundamental reason affecting the washing effect, the process method has always been the focus of research. At present, the most commonly used coal preparation processes at home and abroad are heavy medium coal preparation and flotation. As the flotation technology is only suitable for coal slime with small particle size, its application scope is limited. In contrast, heavy medium coal preparation has the advantages of high separation accuracy and strong adaptability

to raw coal and has gradually become the mainstream coal preparation technology at present [1]. Therefore, this study selects the heavy medium coal preparation plant as the research object and discusses the method to realize multi-objective stable output under the condition of determining the coal preparation process. In addition to process methods, some experts and scholars have studied the methods to ensure coal quality and washout rate from the perspective of management and evaluation. For example, Li [2] put forward 11 suggestions to ensure coal quality, and Gong and Xu [3] studied the performance evaluation method of quality management of coal preparation plant, both of which have played a certain role in promoting the guarantee of coal quality and the improvement of benefits. Other scholars have studied the setting of process parameters in the production process, such as feeding mode and feeding pressure [4–6], established the mapping relationship between process parameters and quality characteristics through historical data, and determined the value range of each parameter. Other scholars have studied the application of quality management tools in the coal preparation plant, such as Elevli [7] studied the application of control chart in the coal preparation plant and realized the abnormal monitoring and control of coal preparation product quality. Xi Jin et al. [8] used the nonlinear regression method to study the optimal combination of various parameters in the coal blending link, which provides a basis for the production control and optimization of coal blending. However, the existing research on coal quality and yield control methods in coal preparation process mainly focuses on postdetection. Even the research on parameter optimization is also the optimization of single-response characteristics, does not consider the common optimization problem between quality and efficiency at the same time, and ignores the influence of noise factors on the robust output of production system in the optimization process. In addition, the existing studies mainly focus on the independent optimization of single factors and lack of systematic research on the impact of coal preparation production system on coal quality and yield. As the heavy medium coal preparation is a huge and complex production system, involving hundreds of equipment and thousands of processes, if the whole system is optimized and analyzed, a large amount of data needs to be collected, even a large number of tests need to be carried out, and the research cost is too high. Therefore, the key subsystems of the whole production system are selected for analysis. The principle of subsystem optimization indirectly reflects the process of realizing the optimal output of the whole production system. As the core-processing process in the production process of heavy medium coal preparation, cyclone separation is the key control link to achieve stable quality and maximum yield of heavy medium coal preparation. It is also the key subsystem in the production system of heavy medium coal preparation. Therefore, the robust parameter design of multiobjective-oriented cyclone production system has become a problem worthy of study. Experts and scholars at home and abroad have done a lot of research on the optimal design of hydrocyclone. Sun et al. [9] and Liu et al. [10], respectively, used the grey correlation method and RSM

(Reynolds stress model) to study the setting of relevant parameters such as volume concentration and flow field of cyclone in heavy medium coal preparation plant. Azizi et al. [11] studied the application of Plackett Burman method in identifying key factors of coal preparation and achieved remarkable results. Chen et al. [12] and Zhang Guo et al. [13] studied the method of parameter control database and linear regression to realize the identification and automatic control of the relationship between parameters s . In addition, many scholars have studied the online optimization and control method of heavy medium coal preparation, the optimal method of coal preparation process design, the optimization method of process yield, and the optimal realization method of cost and quality. [14–20].

However, the aforementioned optimization methods have the following shortcomings: (1) the established optimization function is a function of single-response characteristics, without considering the multiobjective optimization of quality and efficiency; (2) the fitting model established by Plackett Burman experimental design method does not consider the interaction between controllable factors; (3) the existing optimization design methods do not consider the influence of noise on the system output; (4) many existing optimization model data basically come from the historical data of field production, so the representativeness of analysis samples and the effectiveness of data seriously restrict the accuracy of the model. The shortcomings of these studies show that the existing methods cannot determine the value of the best operating parameters that can ensure the robust output of the system under the multiobjective guidance. The study of a system parameter setting is essentially a process of changing the concept of quality management, that is, from the original fire-fighting and passive quality management to preventive and autonomous management. In order to achieve this purpose, many advanced methods have studied by experts and scholars, such as Cheng and Pan [21], Lv et al. [22], and Yan [23], have studied the application of DOE method in the specific production process of machinery manufacturing industry. The significant factors affecting specific output are identified, the optimal parameter setting is confirmed, and the output performance index is greatly improved. Lecureux et al. [24] studied the influence of specific influencing factors on electromagnetic sunscreen through DOE technology, providing data support for the optimization of electromagnetic sunscreen production system. Smith et al. [25] and Moscynski [26] studied the application of DOE technology and method in aerospace field. In terms of precision analysis and control, Yoon et al. [27] applied this method to the parameter optimization analysis of the robot foot control system to realize the optimal parameter configuration of the overall system with multiple inputs and outputs. Doe has also been gradually studied and applied in the fields of chemical industry, service industry, performance evaluation, and cost management [28–30]. For the influence of noise factors on system output, Taguchi [31], a Japanese scholar, first systematically proposed the concept of robust parameter design, which was realized by internal and external table design method and signal-to-noise ratio analysis. Although

this method has been widely used, it is questioned by Miller [32], Montgomery [33], Tsui [34], and other scholars because it does not consider the interaction between factors, the experiment is lack of sequence, and the SNR index is unreasonable. As an improvement, Vining and Myers [35, 36] expressed the RPD problem as a constrained optimization problem and proposed dual-response surface methodology (DRSM) to solve this problem. Its basic idea is to fit the two response surfaces of the process mean and variance respectively, and then minimize the variance with the mean target value as the constraint, so as to improve the robustness of the process. For example, Zhao et al. [37] studied the robust parameter design of double-response surface method in the process of solder paste printing, which improved the quality of solder paste thickness and reduced the volatility. He et al. [38] took the tire product design of a rubber company in Tianjin as the object, optimized an important quality feature (rubber hardness) in the tire production process using the double-response surface method and designed the test design software according to the double-response surface method. The test design results found the ratio conditions to achieve the rubber hardness target and reduce the hardness variation. By adjusting the combination of controllable factors, the double-response surface reduces the sensitivity of the system to noise factors and ensures the robustness of the system. However, the existing double-response surface analysis mainly solves the robust parameter design problem of single-response characteristics, and the optimization of production system includes two dimensions of quality and efficiency and multiple-response characteristics.

Therefore, the main contribution of this research is to explore the method of robust parameter design of hydro-cyclone production system under the condition of multi-response characteristics. First, the double-response surface method for the response characteristics of quality and efficiency is studied, and the double-response function of each response characteristic is established. Then the multiobjective genetic algorithm is used to solve the evaluation problem of multiobjective optimization function under constraints, and finally, the robust parameter values under multiobjective conditions are obtained. The superiority of the proposed method is illustrated by comparing the traditional satisfaction function method with the gradient descent method.

Based on the aforementioned discussions, the main contributions are given as follows:

- (1) In this paper, the dual-response surface method is applied to the robust parameter design of cyclone system for the first time, which provides a new idea for the follow-up research.
- (2) For the parameters that have been estimated, the optimization algorithm is used to optimize, so that more accurate and fine results can be obtained

2. Double-Response Surface Method

Suppose the response variable is Y and the controllable variable is X_1, X_2, \dots, X_k , repeat the measurement on the

response variable Y , and calculate its mean and standard deviation. Taking the mean and standard deviation of the response as output variables, Vining and Myers fit the second-order polynomial model respectively:

$$\begin{aligned}\omega_u(x) &= \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j}^k \beta_{ij} x_i x_j + \varepsilon_u, \\ \omega_\sigma(x) &= \gamma_0 + \sum_{i=1}^k \gamma_i x_i + \sum_{i=1}^k \gamma_{ii} x_i^2 + \sum_{i < j}^k \gamma_{ij} x_i x_j + \varepsilon_\sigma.\end{aligned}\quad (1)$$

As two models are used to optimize the mean and standard deviation, how to select the optimization function to optimize the mean and standard deviation at the same time is a problem that engineers and technicians must solve. For example, when Vining and Myers optimize the quality characteristics of the target type, the optimization method they select is as follows:

$$\left\{ \begin{array}{l} \min \quad \omega_\sigma, \text{ s.t. } \left\{ \begin{array}{l} \omega_u = \tau, \\ L \leq x_i \leq U, \quad i = 1, 2, \dots, k. \end{array} \right. \end{array} \right. \quad (2)$$

where ω is the target value, and l and u are the boundary of x value. Formula (2) can also be transformed into

$$\left\{ \begin{array}{l} \min \quad F(x) = \omega_\sigma^2 + (\omega_u - \tau)^2, \\ \text{s.t.} \quad L \leq x_i \leq U, \quad i = 1, 2, \dots, k. \end{array} \right. \quad (3)$$

3. Multiobjective Genetic Algorithm

The multiobjective optimization problem can be described as follows:

$$\left\{ \begin{array}{l} \min \quad [F_1(x), F_2(x), F_3(x)], \\ \text{s.t.} \quad \left\{ \begin{array}{l} lb \leq x \leq ub, \\ \text{Aeq} * x = \text{beq}, \\ A * x \leq b, \end{array} \right. \end{array} \right. \quad (4)$$

where $\min[F_1(x), F_2(x), \dots, F_m(x)]$ is the objective function to be optimized; X is the variable to be optimized; lb and ub are lower bound and upper bound of variables, respectively; $\text{Aeq} * x = \text{beq}$ is the equality constraint of X ; $A * x \leq b$ and $lb \leq x \leq ub$ are the inequality constraints for X .

Since the objective functions may be contradictory, that is, the improvement of one objective function may require the reduction of another objective function as a cost. As shown in Figure 1, $A1 < B1$ and $A2 > B2$ call such solutions A and B noninferior solutions, or Pareto optimal solutions. The purpose of multiobjective optimization algorithm is to find these optimal solutions.

At present, there are many multiobjective optimization algorithms. Kalyanmoy Deb's fast nondominated sorting genetic algorithm with elite strategy (NSGA-II) is the most widely used one, and this research will adopt a multi-objective optimization algorithm based on NSGA-II.

4. The Double-Response Surface Method Proposed in This Paper

Although the double-response surface method solves the problems of lack of interaction items and lack of sequence of

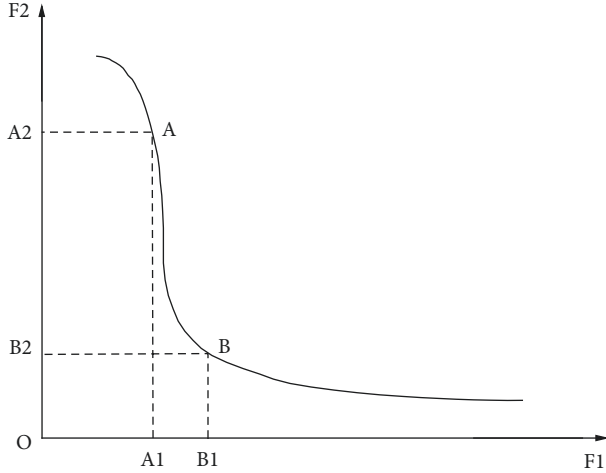


FIGURE 1: Multi objective optimization problem.

experiments in the process of robust design, the current research on double-response surface is mainly to analyze one quality characteristic, and now the products delivered to customers need to be comprehensively evaluated from the perspective of multiple dimensions and multiple quality characteristics. So, the robust parameter design of multi-quality characteristics (response characteristics) has become an urgent problem to be solved, and the multiobjective genetic algorithm just provides the possibility for the realization of robust parameter design of multiresponse characteristics. The traditional optimization methods such as gradient descent method may fall into local extreme values, and the genetic algorithm has a strong ability in the global optimization problem and can solve the problem of local convergence.

For eye-type characteristic Y_1 , large-scale characteristic Y_2 and small-scale characteristic Y_3 , the respective double-response optimization functions are as follows:

$$\begin{aligned} \min F_1(x) &= \omega_{Y_1\sigma}^2 + (\omega_{Y_1u} - \tau)^2, \\ \min F_2(x) &= \omega_{Y_2\sigma}^2 + \left(\frac{1}{\omega_{Y_2u}}\right)^2, \\ \min F_3(x) &= \omega_{Y_3\sigma}^2 + \omega_{Y_3u}^2. \end{aligned} \quad (5)$$

Therefore, the multi response robust parameter optimization function based on the integration of multiobjective genetic algorithm and double response surface method is shown in formula (6).

$$\begin{cases} \min [F_1(x), F_2(x), F_3(x)], \\ \text{s.t.} \begin{cases} lb \leq x \leq ub, \\ Aeq * x = beq, \\ A * x \leq b, \end{cases} \end{cases} \quad (6)$$

Moreover, for the acquisition of test data, the product table is used to design the test, and the response surface test design method is used for the inner table, to ensure that the

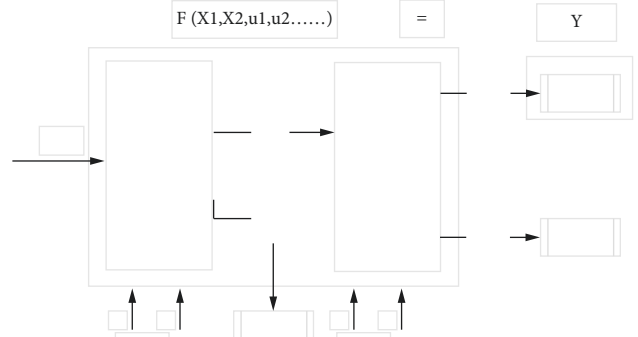


FIGURE 2: Schematic diagram of cyclone production system in a heavy medium coal preparation plant.

established model can reflect the changes of the surface and the influence of interaction terms. According to the two most unfavorable conditions, all the noise factors affecting the inner table test treatment are adopted the comprehensive error method. In other words, each treatment in the internal table runs once under each most unfavorable condition. Therefore, each test treatment will get two values, and the mean and standard deviation of these two values are used as the output response of this treatment.

5. Case Analysis

5.1. Background and Requirements of Test Design. The cyclone production system of a heavy medium coal preparation plant is a pressurized three product heavy medium cyclone production system, which is composed of two product heavy medium cyclone production systems in series. The first section is a high-density area, the products are gangue and medium clean coal, and the second section is a low-density area. Medium clean coal and some small particle size raw coal enter the system for screening, and the products are medium coal and clean coal. The schematic diagram of production process is shown in Figure 2. At present, domestic and foreign coal preparation enterprises adopt the method of “ensuring one end” for the quality control of coal preparation products, focusing on ensuring the quality of cleaned coal products. Coal quality indexes mainly include ash, calorific value, moisture, and so on. Since calorific value, moisture and other indexes can be derived from ash, clean coal ash is selected as the quality response characteristic of this study. In order to maximize the interests of the enterprise, the clean coal washing out rate, that is, the yield index, is also particularly important. Unilaterally meeting the low ash index may meet the requirements of external customers, but the clean coal yield will be greatly reduced, resulting in damage to the interests of the enterprise. Therefore, based on the recent customer demand and underground coal quality, the coal preparation plant decided to set the target of washed coal ash content of the coal preparation plant at 8.50% and the yield target at no less than 69.50% by carrying out the floating and sinking test of raw coal. The floating and sinking test results are shown in Table 1.

TABLE 1: Comprehensive table for float sink test of 50–0.5 mm raw coal.

Density level, kg/L%	Weight, kg	Productivity %			Ash content	Floating object			Cumulative	
		The current level	Total sample	Ash content		The current level, %	Total sample, %	Ash content, %	Productivity, %	Ash content, %
1	2	3	4	5	6	7	8	9	10	
-1.30	0.52	7.40	6.07	3.61	7.40	6.07	3.61	100.00	20.27	
1.30 ~ 1.40	4.41	62.73	51.46	7.19	70.13	57.53	6.81	92.60	21.60	
1.40 ~ 1.50	0.75	10.67	8.75	13.57	80.80	66.28	7.70	29.87	51.85	
1.50 ~ 1.60	0.19	2.70	2.22	22.85	83.50	68.49	8.19	19.20	73.12	
1.60 ~ 1.80	0.09	1.28	1.05	34.01	84.78	69.54	8.58	16.50	81.36	
+1.80	1.07	15.22	12.49	85.34	100.00	82.03	20.27	15.22	85.34	

Note. Sampling location: 301 belts. Total weight of coal sample before floating and sinking: 8.60 kg, ash content: 21%. The red dotted line area focuses on the changes of the two response indicators.

5.2. Test Design Scheme. In order to achieve the set response characteristic y and realize the stability of the system, it is necessary to determine the main controllable factors and noise factors affecting the response characteristics in the cyclone production system, that is, x and u in Figure 1, and reasonably set the value level of each factor in combination with the theory and practical experience of cyclone separation in heavy medium coal preparation plant. After consulting a large number of heavy medium coal preparation data [35], it is known that the cyclone production system includes three parameter types. One is the equipment structure parameters, such as the diameter of feed inlet and discharge outlet, and the other is the operating parameters, such as feed pressure and slime content. The other is fixed parameters, such as the washability of feed coal. Since the structural parameters of the equipment and the mechanism of coal are not determined by the production process, the following part will focus on the robust design of the operating parameters of the cyclone production system under the existing production conditions. Combined with relevant heavy medium coal preparation data, national and industrial standards, and with on-site management, the technicians decided to select the feed pressure (x_1) in the low-density area, the solid-liquid ratio (x_2) in the low-density area, the feed pressure (x_3) in the high-density area and the suspended slurry content in the high-density area (x_4) four parameters are used as important controllable factors, and the value level of each factor is set. Since the density of heavy medium will change in the separation process and cannot be controlled in a short time, two parameters of suspension density (U_1) in low-density area and suspension density (U_2) in high-density area are selected as noise factors, and the upper and lower deviations are allowed to be 5% respectively, as shown in Table 2.

The internal and external table method is used to design the test scheme. The response surface test design method is used to design the 4-factor 2-level test in the inner table. The center points in the test scheme are set as 6 and run 30 times in total. The surface is designed by the most unfavorable comprehensive error method, that is, each treatment in the internal table scheme runs once when the noise factor is the maximum value and the minimum value (one test). Therefore, the robust parameter design of the hydrocyclone

TABLE 2: Controllable factor, noise factor, and factor value level.

Factor	Type	Level
x_1	Controlled	140–160 kPa
x_2	Controlled	1/6–1/4
x_3	Controlled	140–150 kPa
x_4	Controlled	20%–30%
u_1	Uncontrollable	1.38–1.52 kg/cm ³
u_2	Uncontrollable	1.62–1.79 kg/cm ³

production system requires a total of 60 tests. The test arrangement and test results are shown in Table 3, where 1 represents the high-level value, -1 represents the low-level value, and 0 is the central point value.

Neutralization in the table is the test results of ash characteristics under the two most unfavorable conditions, which is the mean value of the two test results, σ_{Y_1} is the standard deviation of the two most unfavorable cases and σ_{Y_2} is the test results and relevant statistical values of yield characteristics.

5.3. Robust Parameter Design of Cyclone Production System

5.3.1. Model Fitting. According to the test design plan and test results, the double-response surface model including quadratic terms for ash characteristics and yield characteristics is established using the double-response surface method, and the effectiveness of the model is analyzed.

Y_1 characteristics, the variance analysis results of the mean obtained by fitting analysis are shown in Table 4, and the estimated effects and coefficients are shown in Table 5.

It can be seen from Tables 4 and 5 that the P of the model is $0.02 < 0.05$, so it is considered that the model is effective, and each estimated effect is significant. It is further verified that the residual error of the model meets the requirements, and there is no abnormality, as shown in Figure 3. Therefore, it can be considered that the fitting model of u_{y1} established is effective.

According to the aforementioned analysis, the fitting model of the mean value of Y_1 characteristic is as follows:

$$U_{Y_1} = 8.4446 + 0.325x_1 - 0.05x_2 + 0.1029x_1^2. \quad (7)$$

TABLE 3: Experimental design scheme and test results.

Serial number	Point type	x_1	x_2	x_3	x_4	Y_{11}	Y_{12}	Y_{21}	Y_{22}	U_{Y1}	U_{Y2}	σ_{Y1}	σ_{Y2}
1	-1	0	0	0	-1	8.45	8.41	69.52	67.23	8.43	68.38	0.028284	1.619270
2	-1	0	0	-1	0	8.50	8.46	69.76	69.55	8.48	69.66	0.028284	0.148490
3	-1	0	-1	0	0	8.52	8.50	69.65	69.56	8.51	69.61	0.014142	0.063640
4	0	0	0	0	0	8.85	8.77	70.12	70.02	8.81	70.07	0.056569	0.070710
5	-1	0	0	0	1	8.50	8.52	69.44	69.57	8.51	69.51	0.014142	0.091920
6	-1	1	0	0	0	8.74	8.70	70.06	69.96	8.72	70.01	0.028284	0.070710
7	-1	0	1	0	0	8.35	8.42	65.44	67.28	8.39	66.36	0.049497	1.301080
8	-1	-1	0	0	0	8.29	8.33	63.25	64.22	8.31	63.74	0.028284	0.685890
9	0	0	0	0	0	8.47	8.32	68.87	64.21	8.40	66.54	0.106066	3.295120
10	-1	0	0	1	0	8.21	8.41	62.76	68.99	8.31	65.88	0.141421	4.405280
11	1	-1	-1	1	1	8.49	8.51	69.51	69.47	8.50	69.49	0.014142	0.028280
12	1	1	-1	-1	1	8.55	8.55	69.66	69.71	8.55	69.69	0.000000	0.035360
13	1	-1	1	-1	1	8.46	8.29	68.89	64.27	8.38	66.58	0.120208	3.266830
14	1	1	-1	-1	-1	8.40	8.51	67.78	69.55	8.46	68.67	0.077782	1.251580
15	1	1	-1	1	1	8.70	8.88	69.86	70.11	8.79	69.99	0.127279	0.176780
16	1	1	1	1	-1	8.50	8.54	69.51	69.62	8.52	69.57	0.028284	0.077780
17	1	-1	-1	1	-1	8.45	8.52	68.88	69.12	8.49	69.00	0.049497	0.169710
18	0	0	0	0	0	8.32	8.41	67.89	68.26	8.37	68.08	0.063640	0.261630
19	0	0	0	0	0	8.33	8.45	67.67	68.23	8.39	67.95	0.084853	0.395980
20	1	-1	-1	-1	1	8.65	8.72	69.99	70.05	8.69	70.02	0.049497	0.042430
21	1	1	1	1	1	8.52	8.54	69.51	69.64	8.53	69.58	0.014142	0.091920
22	0	0	0	0	0	8.35	8.41	67.45	66.53	8.38	66.99	0.042426	0.650540
23	1	-1	-1	-1	-1	8.71	8.72	69.94	69.97	8.72	69.96	0.007071	0.021210
24	1	-1	1	1	-1	8.51	8.55	69.52	69.61	8.53	69.57	0.028284	0.063640
25	1	1	-1	1	-1	8.57	8.67	69.77	69.87	8.62	69.82	0.070711	0.070710
26	1	1	1	-1	-1	8.47	8.54	68.25	69.63	8.51	68.94	0.049497	0.975810
27	0	0	0	0	0	8.32	8.42	65.32	68.71	8.37	67.02	0.070711	2.397090
28	1	1	1	-1	1	8.56	8.50	68.53	69.57	8.53	69.05	0.042426	0.735390
29	1	-1	1	1	1	8.55	8.52	69.55	69.47	8.54	69.51	0.021213	0.056570
30	1	-1	1	-1	-1	8.40	8.60	69.34	69.87	8.50	69.61	0.141421	0.374770

TABLE 4: Estimated effects and coefficients of u_{y1} .

Term	Effect	Coefficient	Coefficient standard error	T	P
Constant		8.4446	0.0344	245.13	<0.001*
A	0.6050	0.325	0.0101	1.02	0.038*
B	-0.1000	-0.0500	0.0181	-1.78	0.047*
A * A	0.2058	0.1029	0.0445	2.31	0.029*

Note.* is significant. S = 0.0119337, R-Sq = 97.48%, R-Sq (adjustment) = 94.11%.

Similarly, the fitting models obtained are as follows:

$$\begin{aligned} \sigma_{Y1} = & 0.05436 - 0.00118x_1 + 0.00471x_2 \\ & - 0.00118x_3 - 0.00432x_4 + 0.0309x_3^2 \\ & - 0.0327x_4^2 - 0.02077x_1x_2 \\ & + 0.01724x_1x_3 - 0.02431x_2x_3, \end{aligned} \quad (8)$$

$$U_{Y2} = 67.657 + 0.43x_1 - 0.45x_2 - 0.05x_4 + 1.615x_4^2,$$

$$\sigma_{Y2} = 1.225 - 0.068x_1 - 0.77x_1^2.$$

According to the double-response surface method in part 2, the optimization functions are constructed for the mean and standard deviation models of Y_1 and Y_2 response characteristics respectively. The optimization function of Y_1 (target type) is shown in formula (9). Since the model established by the model fitting part is obtained based on

coding, the value range of controllable factor is $[-1, 1]$. Since Y_1 is the target characteristic and is strictly not allowed to exceed the set standard by 8.50%.

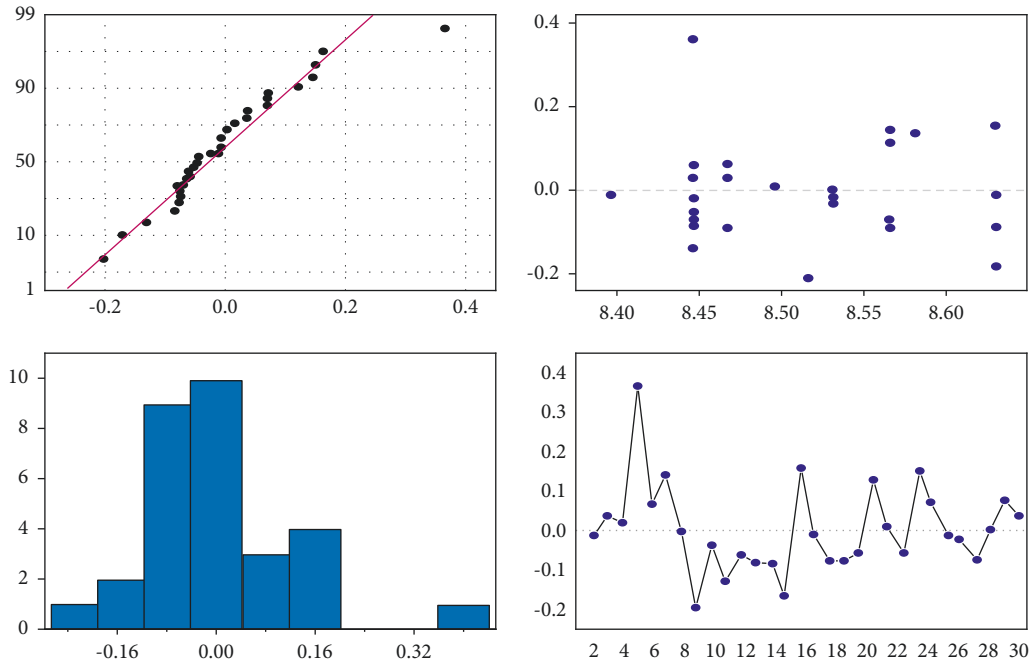
$$\begin{cases} \min & F_1(x) = (U_{Y1} - \tau)^2 + \sigma_{Y1}^2, \\ \text{s.t.} & \begin{cases} -1 \leq x_i \leq 1, i = 1, 2, 3, 4 \\ U_{Y1} \leq 8.50\% \end{cases} \end{cases} \quad (9)$$

Similarly, the optimization function of Y_2 characteristic (expected to be large) can be obtained as follows:

$$\begin{aligned} \min F_2(x) = & \left(\frac{1}{U_{Y2}} \right)^2 + \sigma_{Y2}^2 \\ \begin{cases} -1 \leq x_i \leq 1, & i = 1, 2, 3, 4 \\ U_{Y2} \geq 69.5\% \end{cases} \end{aligned} \quad (10)$$

TABLE 5: Analysis of variance of u_{y1} fitting model.

Source	Freedom	Seq SS	Adj SS	Adj MS	F	P
Model	3	0.399350	0.399350	0.049919	88.09	0.002
Linear	2	0.226562	0.001765	0.107832	1.78	0.045
A	1	0.003306	0.003306	0.003306	5.83	0.045
B	1	0.223256	0.223256	0.223256	393.98	<0.001
Square	1	0.139194	0.139194	0.019885	35.09	0.007
A * A	1	0.139194	0.139194	0.019885	35.09	0.007
Error	3	0.001531	0.001531	0.001531	2.70	0.199
Misfit	1	0.001700	0.001700	0.000567	0.064	0.001
Pure error	2	0.001700	0.001700	0.000567		

FIGURE 3: Residual plot of U_{Y1} model.

According to the robust parameter design method integrating double-response surface and multiobjective genetic algorithm proposed in the third part of the paper, the optimization function of multiresponse characteristics can be obtained.

The number is as follows:

$$\begin{aligned}
 \min F_1(x) &= (U_{Y1} - \tau)^2 + \sigma_{Y1}^2, \\
 \min F_2(x) &= \left(\frac{1}{U_{Y2}}\right)^2 + \sigma_{Y2}^2, \\
 &\begin{cases} -1 \leq x_i \leq 1, i = 1, 2, 3, 4 \\ U_{Y1} \leq 8.50\% \\ U_{Y2} \geq 69.5\% \end{cases} .
 \end{aligned} \tag{11}$$

Since the optimization function (11) includes nonlinear constraints, the external penalty function method is used to transfer the nonlinear inequality constraints to the objective

function to obtain the final optimization function, as shown in formula (12), where m is a very large penalty coefficient, which is also the fitness function expression of multiobjective genetic algorithm.

$$\begin{aligned}
 \min F_1(x) &= (U_{Y1} - \tau)^2 + \sigma_{Y1}^2 \\
 &\quad + M * \min((8.50\% - U_{Y1}), 0)^2, \\
 \min F_2(x) &= (1/U_{Y2})^2 + \sigma_{Y2}^2 \\
 &\quad + M * \min((U_{Y2} - 69.5\%), 0)^2, \\
 \text{s.t. } &-1 \leq x_i \leq 1, \quad i = 1, 2, 3, 4.
 \end{aligned} \tag{12}$$

The parameters of the multiobjective genetic algorithm are set as follows: the optimal front-end individual coefficient Pareto fraction: 0.3, population size: 100, maximum iteration generations: 200, stop algebra limit: 200, fitness function value deviation function: $1e - 100$. The obtained Pareto front-end distribution and the distance between individuals in the evolution process are shown in Figure 4.

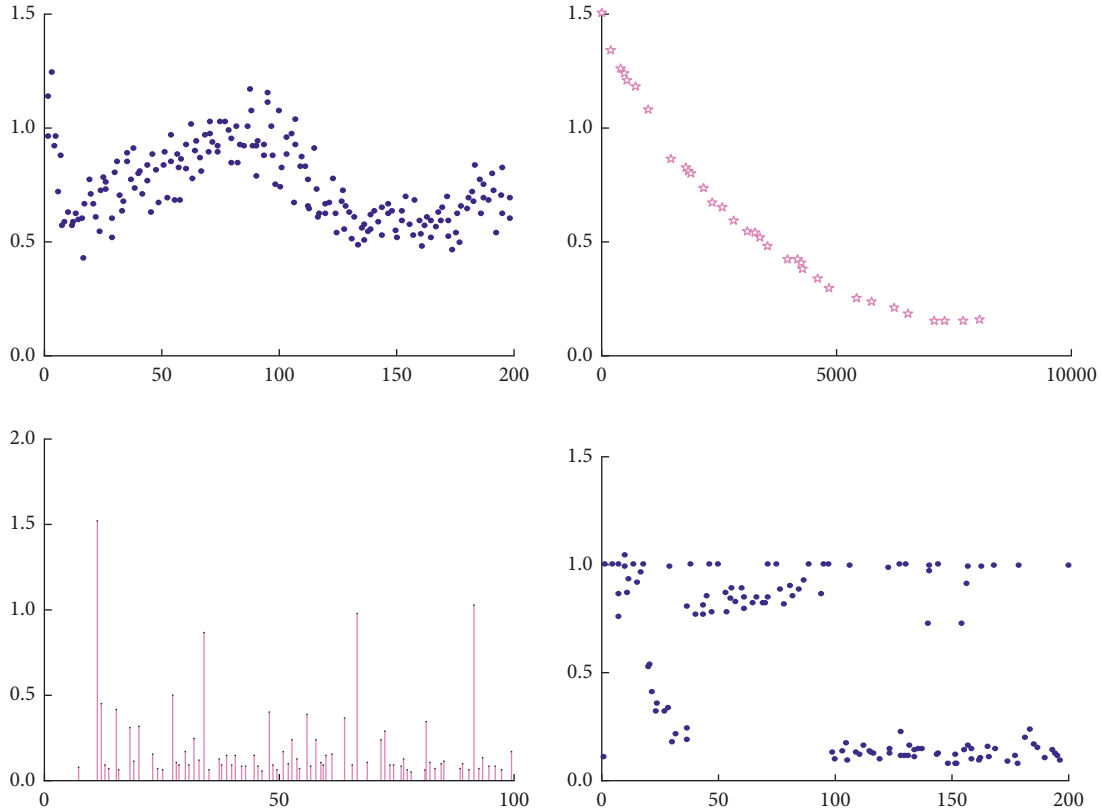


FIGURE 4: Analysis diagram of multiobjective genetic algorithm.

According to the Pareto front-end distribution in Figure 4, the junction distribution is relatively uniform, indicating that the obtained solution is effective and the effect is good. Meanwhile, the other three subgraphs in the figure show that the distance and propagation speed between individuals are gradually decreasing, which indicates that with the increase of evolution, the diversity of the population gradually decreases and the solution gradually converges. Since the optimal front-end individual coefficient is 0.3 and the number of populations is 100, 30 groups of optimal solutions are obtained, of which 4 groups are close to or meet the requirements, as shown in Table 6.

As the Y_1 characteristic is required to be strictly less than or equal to 8.50%, only the third group solution in Table 6 strictly meets the requirements. Therefore, the robust parameter design value (code value) of the hydrocyclone production system is (X_1 : 0.050332, X_2 : -0.772998, X_3 : -0.295515, X_4 : 0.999956), and the actual values are 150.50 kpa, 0.176143.52 kpa, and 30%, respectively. The optimal response characteristics that can be realized are Y_1 : 8.499% and Y_2 : 69.59%.

5.4. Comparative Analysis of the Effectiveness of the Method. The optimization function of each output characteristic established by the double-response surface method actually adopts the method of transforming constraints into objective functions, so as to transform them into unconstrained single-objective optimization problems. At present, most of

the optimization functions with multiple output characteristics are transformed by the satisfaction function method. After transforming the multiobjective function into a single-objective function, the stability point of the system is obtained by derivation, the type of stability point is determined by canonical analysis, and the confidence interval of the optimal value is estimated according to the normal hypothesis. Therefore, this part shows the effectiveness and efficiency of this method by comparing the differences between the traditional optimization method and the proposed optimization method.

First, if each output characteristic optimizes its dual-response optimization function respectively, set the upper limit of u_{y_1} to 8.500001%, the target to 8.50%, and the lower limit to 8.40%. The type is the desired type, and set σ_{Y_1} is expected to be large. The response optimizer of Minitab is used to optimize the two functions at the same time. The results are shown in Figure 5.

Similarly, the lower limit of u_{y_2} is set to 69.5%. In order to achieve large scale, the double-response optimization result of Y_2 is obtained, as shown in Figure 6.

Considering two characteristics and four response functions at the same time, the results are shown in Figure 7, It can be seen from the figure that the distribution of two characteristics and four response function is relatively uniform, thus demonstrating the effectiveness of the proposed method. The optimization results of the proposed method and the aforementioned analysis results are summarized and Table 7 is obtained. It can be seen from the table

TABLE 6: Partial optimal solutions obtained by multiobjective genetic algorithm.

x_1	x_2	x_3	x_4	U_{Y1}	σ_{Y1}	U_{Y2}	σ_{Y2}
0.299078	-0.414492	0.008521	0.998234	8.571729	0.017855	69.531513	1.135788
0.262752	-0.621845	-0.235414	0.998063	8.568191	0.014995	69.608660	1.153973
0.050332	-0.772998	-0.295515	0.999956	8.499869	0.011689	69.591351	1.219627
0.352806	-0.421071	-0.274353	0.997759	8.593124	0.016355	69.556071	1.105166

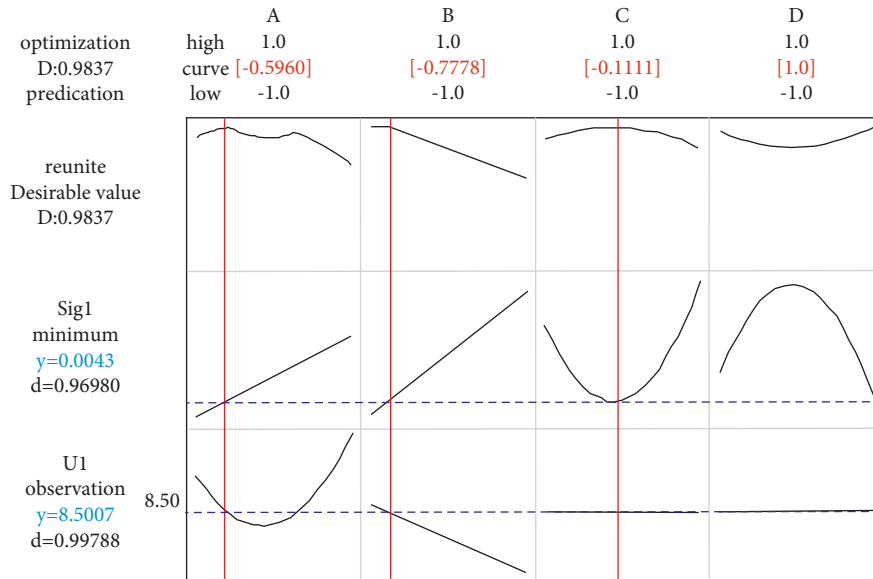


FIGURE 5: Double-response optimization analysis diagram of Y_1 characteristics.

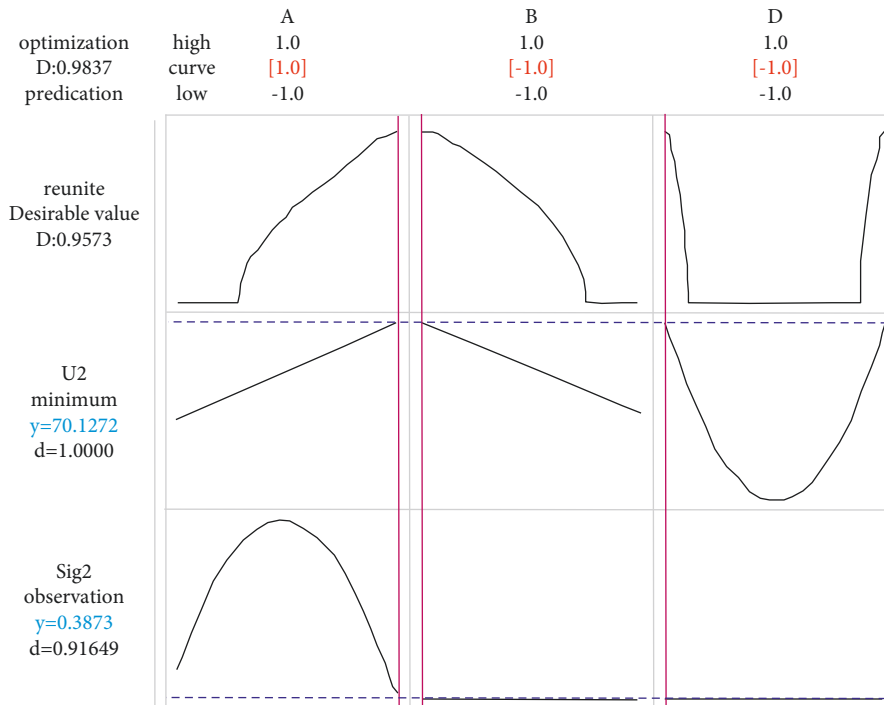


FIGURE 6: Double-response optimization analysis of Y_2 characteristics.

that better output results can be obtained when a single output characteristic is optimized respectively. However, the calculated parameter value (factor values) are quite different,

indicating that their respective optimizations cannot guarantee the overall optimization of the system. When considering two output characteristics at the same time, the

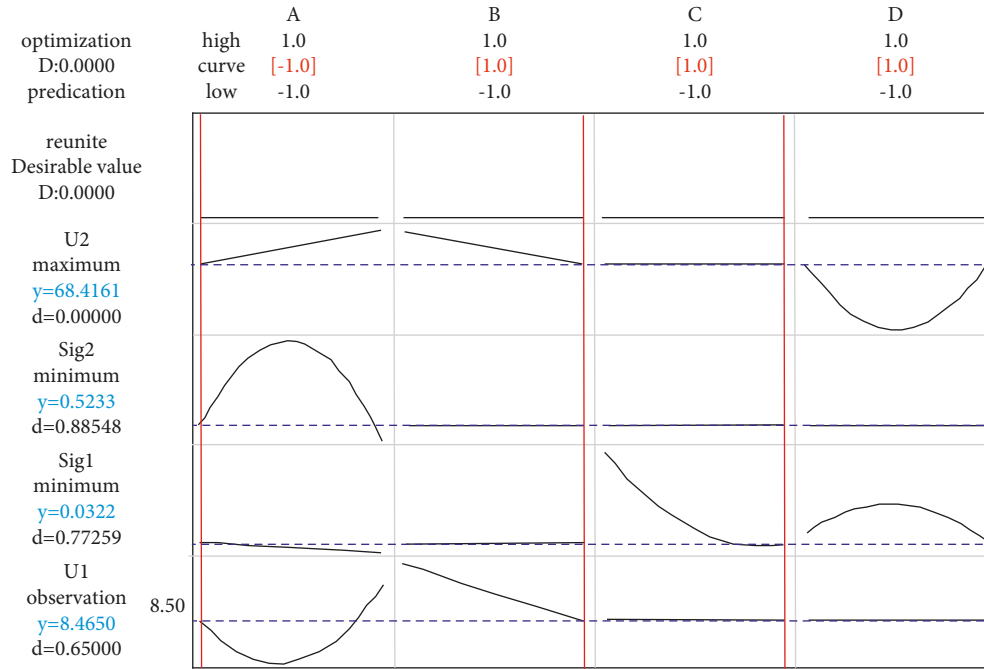
FIGURE 7: Optimization analysis of Y_1 and Y_2 response characteristics.

TABLE 7: Comparative analysis.

Methods	x_1	x_2	x_3	x_4	U_{Y_1}	σ_{Y_1}	U_{Y_2}	σ_{Y_2}
Sole Y_1	-0.5960	-0.7778	-0.1111	1.0	8.5007	0.0043	—	—
Sole Y_2	1	-1	0	-1	—	—	70.1273	0.3873
Sole Y_1, Y_2	-1	1	1	1	8.4650	0.0322	68.4161	0.5233
Proposed method	0.050332	-0.772998	-0.295515	0.999956	8.499869	0.011689	69.591351	1.219627

traditional method is used to optimize four response functions at the same time. Although the standard deviation of Y_2 characteristics is twice as small as that of the proposed method, the performance of the traditional method is not as good as that of the proposed method in the other three indexes, especially for Y_1 ash characteristics, the optimization results are not only closer to the ash requirement of 8.50%, but also have less fluctuation, so as to ensure that the coal preparation plant has high process control ability. Moreover, the maximum yield obtained by the traditional method is 68.4146%, which does not meet the design requirement of 69.50% in the floating and sinking test. In conclusion, it can be seen that the method proposed in this study is better than the traditional multiresponse robust parameter design method as a whole.

6. Conclusion

This paper shows the feasibility of integrating double-response surface method and multiobjective genetic algorithm to solve the problem of robust parameter design under the condition of multiresponse characteristics of cyclone production system in heavy medium coal preparation plant is discussed. The conclusions are as follows: (1) the optimal value of response characteristic ash is 8.499%, the optimal value of yield is 69.59%, and the

optimal controllable factors are 150.50 kpa, 0.176143.52 kpa, and 30%; (2) A double-response optimization model between response characteristics and influencing factors is established; (3) The effectiveness of double-response surface method and multiobjective genetic algorithm in the optimization process of production system of heavy medium coal preparation plant is verified, and the qualified and stable coal quality and maximum yield are realized; (4) It has changed the fire-fighting and passive quality management mode of coal preparation enterprises, and gradually changed to the preventive, autonomous, and stable quality management mode. The enterprises are no longer tired of dealing with the continuous coal quality problems, but realize the expected output by controlling the stability of influencing factors. When the whole system is gradually stable, the coal quality inspection link of coal preparation plant can even be cancelled or weakened; (5) Coal preparation enterprises are more confident to realize the accurate control of the production process. In conclusion, the proposed method can explore the relationship between system input and output, reduce the influence of noise on model output and determine the robust parameter values under the condition of multiresponse characteristics, thus providing scientific support for the transformation of quality management mode.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] F Liu, "Research and development of coal preparation technology with heavy medium cyclone," *Coal preparation technology*, vol. 05, pp. 1–13, 2006.
- [2] C. Li, "Zhou Wei Ensure the quality of cleaned coal and improve the yield of cleaned coal," *Coal Technology*, vol. 03, pp. 96–97, 2009.
- [3] D. Gong and S. Xu, "Study on quality management performance assessment of coal preparation plant," in *Proceedings of the Computer Applications in the Minerals Industries*, pp. 495–498, Shers, Beijing, China, November 2020.
- [4] S Yue, "Reforming the feeding mode to improve the stability of clean coal quality," *Coal quality technology*, vol. 01, pp. 27–28, 2004.
- [5] J Liu, "Separation effect and influencing factors of non-pressure feeding three product heavy medium cyclone in Fangezhuang Coal Preparation Plant," *Clean coal technology*, vol. 04, pp. 10–11, 2010.
- [6] L. Sun, "Kuang Yali Design of real-time prediction system platform for coal preparation production process parameters," *Mining machinery*, vol. 01, pp. 85–89, 2012.
- [7] S. Eleveli, "Coal quality control with control charts," *Coal Preparation*, vol. 26, no. 4, pp. 181–199, 2006.
- [8] G. Xi-Jin, C. Ming, and W. Jia-Wei, "Coal blending optimization of coal preparation production process based on improved GA," *Procedia Earth and Planetary Science*, vol. 1, no. 1, pp. 654–660, 2009.
- [9] W. Sun, L. Shen, and J. Chen, "Shi Shangjie Application of grey correlation in main cause analysis of separation effect of heavy medium cyclone," *Mining machinery*, vol. 06, pp. 75–80, 2013.
- [10] F. Liu, A. Qian, and X. Guo, "Selection of turbulence numerical calculation model for flow field of heavy medium cyclone," *Journal of coal*, vol. 03, pp. 346–350, 2006.
- [11] D. Azizi, M. Gharabaghi, and N. Saeedi, "Optimization of the coal flotation procedure using the Plackett-Burman design methodology and kinetic analysis," *Fuel Processing Technology*, vol. 128, pp. 111–118, 2014.
- [12] J. Chen, K. W. Chu, R. P. Zou et al., "How to optimize design and operation of dense medium cyclones in coal preparation," *Minerals Engineering*, vol. 62, pp. 55–65, 2014.
- [13] W. Zhang-guo, K. Ya-li, L. Zhe, and S. Chang-sheng, "Model of coal product structure based on particle swarm optimization algorithm," *Procedia Earth and Planetary Science*, vol. 1, no. 1, pp. 640–647, 2009.
- [14] T. M. Romberg, N. R. Godfrey, R. J. Stevens, and M. R. Davies, "On-line optimization and control of heavy media circuits in coal preparation processes," *National Conference Publication - Institution of Engineers*, vol. 90, no. 6, pp. 1–5, 1990.
- [15] V. Gupta and M. K. Mohanty, "Coal preparation plant optimization: a critical review of the existing methods," *International Journal of Mineral Processing*, vol. 79, no. 1, pp. 7–17, 2006.
- [16] A. Vishal Mahajan, M. K. Mohanty, and S. K. Biswal, "Performance optimization of a coal preparation plant using Genetic Algorithms," *American Society of Mechanical Engineers, Dynamic Systems and Control Division (Publication) DSC*, vol. 73, no. 1A, pp. 717–726, 2004.
- [17] V. Gupta, M. Mohanty, and A. Mahajan, "Genetic algorithms - a novel technique to optimize coal preparation plants," *International Journal of Mineral Processing*, vol. 84, no. 1–4, pp. 133–143, 2007.
- [18] G. H. Luttrell and R. Q. Honaker, "Practical optimization of coal preparation plants," in *Proceedings of the Jan D. Miller Symposium - Innovations in Natural Resource Processing*, pp. 367–382, Salt Lake City, CA, USA, March 2005.
- [19] J. Li, K. Yali, W. Zhanguo, D. Bo, L. Yunyu, and C. Bing, "The application of a differential evolution algorithm on products structure optimization in coal preparation plants," *International Journal of Coal Preparation and Utilization*, vol. 32, no. 1, pp. 1–10, 2012.
- [20] C. Wang, J. Hu, X. He, and Q. Ren, "Study on the preparation optimization and characteristics of modified media for coal-mine wastewater treatment with high iron and high manganese concentration," in *Proceedings of the 2011 International Symposium on Water Resource and Environmental Protection*, pp. 1162–1165, IEEE Computer Society, Xi'an, China, May 2011.
- [21] L. Cheng and X. pan, "Yan Qingdong, etc Influence and optimization of blade number of hydraulic torque converter based on DOE and RSM," *Journal of Beijing University of Technology*, vol. 32, no. 7, pp. 689–693, 2012.
- [22] J. Lv, J. Chen, and Y. Zheng, "Etc Structural optimization design of excavator exhaust muffler based on DOE," *Noise and vibration control*, vol. 2014, no. 5, pp. 210–213, 2014.
- [23] Li Yan, "Cheng Junkang Application of Six Sigma DOE test method in motor production improvement process," *Micro*, vol. 47, no. 7, pp. 89–91, 2014.
- [24] M. Lécureux, C. Deumié, S. Enoch, and M. Sergent, "Electromagnetic sunscreen model: design of experiments on particle specifications," *Applied Optics*, vol. 54, no. 28, pp. 8369–8374, 2015.
- [25] N. Smith, J. Camberos, and E. Alyanak, "Design of experiments for aeroelastic analysis," in *Proceedings of the ASME International Mechanical Engineering Congress and Exposition, Proceedings*, pp. 49–57, American Society of Mechanical Engineers (ASME), Houston, TX, USA, November 2012.
- [26] M. J. Moscynski, "Design of experiments for aerospace products," *SAE Technical Papers*, vol. 2009, no. 01, p. 3138, 2009.
- [27] J. Yoon, T. Park, and K. Jun, "Optimum design of the humanoid robot's foot using the design of experiments," in *Proceedings of the IASTED International Conference on Telematics[C]. Würzburg, Germany: International Association Science and Technology for Development (IASTED); Technical Committee on Robotics; Technical Committee on Telecommunications*, pp. 431–436, Würzburg, Germany, August 2007.
- [28] B. Nabil, A. Aissa, and B. I. Aguida, "Use of a new approach (design of experiments method) to study different procedures to avoid plastic shrinkage cracking of concrete in hot climates," *Journal of Advanced Concrete Technology*, vol. 9, no. 2, pp. 149–157, 2011.
- [29] M. H Wan, H. Wan, M. Nizam, R Abdul, and A. G. Jaharah, "Planning production and control: design of experiments (DOE) in legal service industry," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 04, no. 12, pp. 1684–1689, 2012.
- [30] P. Mallick, S. W. Meador, R. F. Shangraw, and R. John, "DOE benchmarking for cost management," in *Proceedings of the*

- AACE International. Transactions of the Annual Meeting*, AACE, San Francisco, CA, USA, June 1994.
- [31] G. Taguchi, "The system of experimental design engineering methods to optimize quality and minimize cost[M]," *American Supplier Institute*, vol. 1, no. 1, pp. 7-8, 1987.
 - [32] A. Miller, "Analysis of parameter design experiments for signal-response systems," *Journal of Quality Technology*, vol. 34, no. 2, pp. 139-151, 2002.
 - [33] D. C. Montgomery, *Design and Analysis of Experiments*, Wiley, Hoboken, NJ, USA, 2007.
 - [34] K.-L. Tsui, "A critical look at Taguchi's modelling approach for robust design," *Journal of Applied Statistics*, vol. 23, no. 1, pp. 81-96, 1996.
 - [35] R. H. Myers and W. H. Carter, "Response surface techniques for dual response systems," *Technometrics*, vol. 15, no. 2, pp. 301-317, 1973.
 - [36] G. G. Vining and R. H. Myers, "Combining Taguchi and response surface philosophies: a dual response approach," *Journal of Quality Technology*, vol. 22, no. 1, pp. 38-45, 1990.
 - [37] M. Zhao, E. pan, and Y Guo, "Robust parameter design based on double response surface method," *Industrial Engineering & Management*, vol. 15, no. 1, pp. 87-91, 2010.
 - [38] Z. He, S. Zhang, and Z. Liu, "Application of double response surface method in improving product design," *System engineering theory and practice*, vol. 21, no. 9, pp. 140-144, 2001.

Research Article

Head CT Image Segmentation and Three-Dimensional Reconstruction Technology Based on Human Anatomy

Zhenyu Wu ¹, Lin Wang ¹, Yifei Li ¹, Shuhui Dai ², and Dongliang Zhang ³

¹Department of Anatomy, Histology and Embryology and K.K. Leung Brain Research Centre, The Fourth Military Medical University, Xi'an 710032, China

²Department of Neurosurgery, Xijing Hospital, Fourth Military Medical University, Xi'an 710032, China

³Department of Burns and Cutaneous Surgery, Xijing Hospital, The Fourth Military Medical University, Xi'an 710032, China

Correspondence should be addressed to Shuhui Dai; weijl@fmmu.edu.cn and Dongliang Zhang; zhangdl@fmmu.edu.cn

Received 24 March 2022; Revised 29 April 2022; Accepted 18 May 2022; Published 16 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Zhenyu 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.

With the continuous development of computer science and technology, the level of medical image processing and analysis technology has been significantly improved. In order to further optimize the medical imaging technology and provide assistance for medical diagnosis and treatment, this study will explore the head CT image segmentation technology and three-dimensional reconstruction technology based on human anatomy, using two morphological operation methods of image expansion and image corrosion, as well as the triangulation method based on surface contour. Optimize CT image segmentation technology and three-dimensional reconstruction technology. The results show that the CT image segmentation technology based on human anatomy can obtain the more essential morphology and features of the target image, and significantly improve the image quality. The size of the threshold can have a certain impact on the 3D reconstruction effect and reconstruction time to a certain extent. The larger the threshold, the shorter the reconstruction time, but the worse the 3D reconstruction effect. This shows that the target image after fitting has a good reconstruction effect, but the threshold level should be kept at a low level. The head CT image segmentation technology and three-dimensional reconstruction technology based on human anatomy have good application effects and can be popularized and applied in clinical diagnosis and treatment.

1. Introduction

In the process of the continuous development of medical technology, the application scope of human anatomy has been significantly expanded. Applying it to the image segmentation of head computed tomography (CT) can improve the effect of medical image segmentation to a certain extent [1]. The application of medical image processing and analysis in clinical medicine is extremely critical and has attracted more and more attention. In order to improve the accuracy and reliability of medical diagnosis, it is imperative to explore and optimize medical image segmentation and processing technology and three-dimensional reconstruction technology [2]. CT image segmentation and processing technology plays an important role in the diagnosis and detection of head tumors. It can effectively realize image

segmentation and processing when the boundary between normal head and diseased soft tissue is poor [3]. Three-dimensional reconstruction technology is a visualization means of scientific calculation, which can effectively analyze the calculation results of computer data and understand the changes of data in the calculation process under the action of computer graphics and image processing technology [4]. In view of this, this study will make a detailed exploration on the head CT image segmentation and three-dimensional reconstruction technology based on human anatomy, in order to improve the effectiveness of image segmentation and processing technology and three-dimensional reconstruction technology and promote the sustainable development of medical image processing technology.

This study mainly discusses the mechanism and application effect of head CT image segmentation technology and

three-dimensional reconstruction technology in the context of human anatomy. Morphological operation is mainly used to explore the segmentation technology of head CT image. Morphological operation can effectively repair the area of the target CT image. Under the action of image expansion and image corrosion, head CT image can be well segmented. In addition, the surface rendering method in three-dimensional reconstruction technology is also analyzed in detail. This method is a triangulation method based on surface contour, which can use the small plane with the shape of triangle or polygon to place it in the boundary contour line as a filler to form a complete and smooth object surface. The combination of the two research methods provides strong support for exploring the application effect of head CT image segmentation and three-dimensional reconstruction technology.

In this study, the most innovative point is that in morphological operation, two operation methods of image expansion and image corrosion are used at the same time, which can effectively deal with the diversified problems in image segmentation. This innovative method can separate two connected objects on the basis of filling the hole after image segmentation so that the target image is not negatively affected by its own internal segmentation or external object adhesion, which can significantly improve the technical level of CT image segmentation and obtain more objective and accurate image segmentation results. Moreover, the research also innovatively makes a comprehensive comparative analysis of the threshold, three-dimensional reconstruction effect, and reconstruction time, which makes the application process of three-dimensional reconstruction technology not only blindly improve the threshold level but also comprehensively consider the impact of reconstruction efficiency and reconstruction effect.

The main structure of the paper is divided into four parts. The core content of the second part is to explore the research progress of many experts and scholars at home and abroad in CT image segmentation technology and three-dimensional reconstruction technology and briefly discuss its research means and research results. The third part mainly probes into the action mechanism of the two core technologies in this research, including the binarization processing of head CT image segmentation technology, the repair and merging of image regions, the construction of the contour model, the triangulation surface drawing method of three-dimensional reconstruction technology, and the simplification processing method of the triangular mesh. The fourth part focuses on the analysis of the application effect of head CT image segmentation technology and three-dimensional reconstruction technology. The fifth part mainly expounds all the research work and research results, summarizes the disadvantages and deficiencies, and looks forward to the future research direction.

2. Related Work

In recent years, medical diagnosis and treatment methods have been fully developed in the process of clinical application of medical imaging technology. However, in order to

prevent the distortion of lesion target location and improve the effectiveness of medical diagnosis and treatment, the research on medical image processing technology at home and abroad has been continuously promoted and gradually deepened. Mason et al. believe that the current treatment workflow based on cone beam computed tomography (CBCT) is easily limited by the contrast difference of soft tissue. Therefore, ultrasonic fusion CBCT is used to compare and analyze the consistency and confidence of images. The accuracy of image segmentation and location is successfully improved [5]. Aiming at the problem of great difficulty and variability in liver segmentation, researchers such as Mourya et al. used CT image segmentation technology to eliminate the lack of clear edge of the liver boundary, significantly estimated the area of liver tumor, and carried out accurate automatic segmentation [6]. Selvaraj et al. used CT for image scanning and quantitative processing to solve the problem of Covid-19 area contour description and successfully realized the effective segmentation of Covid-19 infection severity [7]. Zeng et al. found that it is difficult to realize the automatic segmentation of head image from ultrasonic image and head circumference biometrics. By using DAG v-net deep learning model, CT technology was used to segment fetal ultrasonic image effectively [8]. Eckl et al. believe that image-guided radiotherapy can benefit from the implementation of radiotherapy technology. They have completed the construction and training of body part-specific model by evaluating the conversion algorithm from cone beam computed tomography to synthetic CT based on cyclic generation countermeasure network. Finally, CT images with high quality were generated [9]. Tappeiner et al. found that in the process of radiotherapy for head and neck cancer, the focus is to accurately depict the organs with risk in the image. The study used the method of deep learning combined with CT to effectively train the data set and successfully obtained the segmented image with high accuracy [10].

Yamamoto et al. conducted in-depth research on osteoclasts, used continuous section scanning electron microscope technology and three-dimensional reconstruction technology to reconstruct osteoclasts, and successfully found the unique three-dimensional shape of Golgi [11]. Yushan et al. analyzed rare congenital malformations and their related characteristics, achieved surgical reduction by using image technologies such as three-dimensional reconstruction, and finally achieved a better effect of surgical intervention [12]. In order to improve the accuracy and reliability of 3D reconstruction, Xie et al. applied the multiview photometric stereo fusion algorithm to 3D reconstruction technology and verified it through multiple iterations. Finally, the 3D reconstruction results with higher precision surface normals and better quality were successfully obtained [13]. Wang et al. believed that the two-dimensional image has the characteristics of fast but simple, so they applied the mobile robot vision technology to it, improved it, proposed the three-dimensional reconstruction technology, and verified the efficiency and accuracy of the three-dimensional reconstruction technology through simulation analysis [14]. Ruiz et al. proposed a nonrigid object 3D

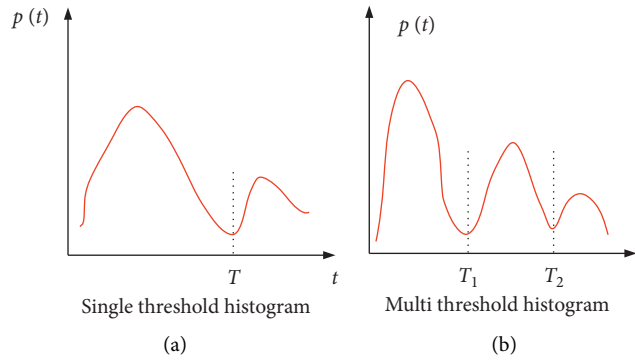


FIGURE 1: Different types of threshold histograms. (a) Single-threshold histogram; (b) multithreshold histogram.

reconstruction algorithm using depth camera, which can dynamically construct the dense 3D model of the target object, improve the accuracy of 3D reconstruction, and can be applied to many fields such as medicine, agriculture, and human-computer interaction [15]. Jani et al. found that the application of three-dimensional technology in the reconstruction and description of human bone fragments is less. By analyzing the application of three-dimensional model in forensic cases, the accuracy of three-dimensional reconstruction model is improved from the perspective of anatomical features and digital analysis, and the optimization of three-dimensional reconstruction technology is realized [16].

By analyzing the research results of many scholars at home and abroad in CT image segmentation and three-dimensional reconstruction technology, it can be found that these two technologies have made great progress and have a wide range of applications, especially medical diagnosis and treatment. CT image segmentation technology and three-dimensional reconstruction technology can reflect a good application effect in medical diagnosis and more intuitively and accurately present the image information of the focus, which is convenient for later clinical treatment intervention. In view of this, this study takes human anatomy as the basic basis to deeply explore CT image segmentation technology and three-dimensional reconstruction technology in order to improve the accuracy of medical image segmentation and three-dimensional reconstruction.

3. Research on CT Image Segmentation Technology and Three-Dimensional Reconstruction Technology

3.1. Patching and Merging of Image Regions and Construction of Contour Model. In the actual process of computed tomography, X-rays are projected and scanned at different angles only in specific faults, and then, the corresponding projection and scanning results are handed over to the computer for further processing so as to obtain the two-dimensional projection distribution results of specific faults in the object to be measured. In the context of human anatomy, there are some differences between various human tissues and other objects. Taking the head as an example,

when forming the CT image of the head, the X-ray absorption of the head tissue is different, so the CT image segmentation technology segments a section of the head tissue, and the small cube obtained after segmentation is voxel. When X-ray passes through the head, the density or gray level in each single voxel is measured as a pixel, that is, the basic unit in CT image. Before applying CT image segmentation technology, it is necessary to binarize the image. Usually, threshold segmentation based on image histogram is adopted as shown in Figure 1.

Figure 1(a) is a single-threshold histogram. Its segmentation principle is to use a threshold T to classify the gray level of the target image, determine the pixels whose gray value is in the same gray range as the segmented image from the same target, and finally realize accurate segmentation of the target image. The segmentation principle of the multithreshold histogram shown in Figure 2(a) is basically consistent with that of the single-threshold histogram. The difference is that more than one threshold is used, that is, T_1 and T_2 . The threshold processing can be regarded as the test of function T , see (1) for details.

$$T = T[x, y, p(x, y), f(x, y)]. \quad (1)$$

In (1), T represents the threshold, (x, y) represents a point in the image, $f(x, y)$ represents the gray level corresponding to the point, and $p(x, y)$ represents its local properties. If the point is taken as the center to obtain the average gray level of the neighborhood, the image after threshold processing can be defined as given in the following equation:

$$g(x, y) = \begin{cases} 1, & f(x, y) > T, \\ 0, & f(x, y) \leq T. \end{cases} \quad (2)$$

In equation (2), $g(x, y)$ represents an image after threshold processing. According to equation (2), the pixels marked 1 and 0 correspond to the target object and the background, respectively. After the binarization of the image is completed, the region can be repaired and merged by morphological operation. The most commonly used operations of region repair are image expansion and image corrosion. The former mainly refers to expanding the boundary points of binary objects and integrating the background points in contact with them so as to achieve the

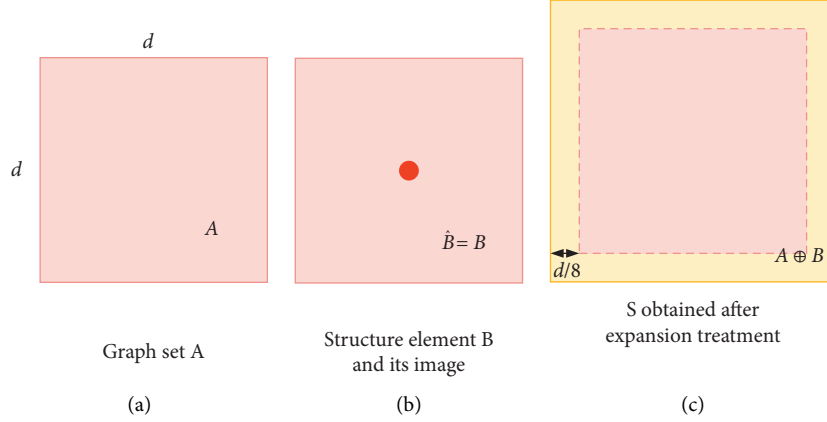


FIGURE 2: Schematic diagram of image expansion process. (a) Graph set A ; (b) structure element B and its image; (c) S obtained after expansion treatment.

expansion effect of image expansion to the outside. This operation method is usually applied to filling the holes after image segmentation as shown in the following formula:

$$S = A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\}. \quad (3)$$

S represents the expansion result of the binary image set; B represents the structural element used for expansion, and its value is 0 or 1. Equation (3) shows that the S is obtained by using the graph B to expand the graph A . The process diagram of expansion treatment is shown in Figure 2.

Figure 2(a) shows a simple graph set A ; Figure 2(b) shows the structural element B and its image, which have a symmetrical and equal relationship with respect to the central point; Figure 2(c) shows the S obtained after A expansion, from which the expansion range can be clearly seen. Another operation of region repair is image corrosion. The action mechanism of this method is quite different from image expansion. Its principle is to eliminate the boundary points of objects and shrink them inward. It is usually applied to the separate processing of two connected objects. Image expansion and image corrosion are both effective ways to trim the target object. After local or overall processing, the corresponding change characteristics can be obtained. In order to maintain the main features and shape of the head contour, it is usually necessary to be careful when selecting structural elements and avoid being too large or too small. After the binary processing and morphological processing of the image, the target image will form multiple subregions, which need to be segmented and extracted by region growth algorithm so as to obtain the initial contour of the head. The last step of CT image segmentation technology is to construct the contour model, and the most common method is to construct the set active contour model through the level set method. Different from other solving methods of geometric curve evolution, the level set method does not directly track the motion of evolution curve but uses implicit expression to describe the plane closed curve. First, let the mathematical expression of the plane active curve be as given in the following equation:

$$C(p, t) = (x(p, t), y(p, t)). \quad (4)$$

In equation (4), p represents any parametric variable and t represents time variable, then the evolution process of curve C in the direction of unit normal vector \vec{N} is shown in the following equation:

$$\frac{\partial C}{\partial t} = V(C) \vec{N}. \quad (5)$$

The curve evolution in equation (5) mainly includes two different forms of constant evolution and curvature evolution. In (5), it represents a velocity function that determines the velocity of each point on the active curve. The evolution of curves mainly includes two different forms of constant evolution and curvature evolution. The difference between them lies in the setting of velocity function, i.e., constant and curvature. The evolution of the former can be expressed as a partial differential equation shown in the following equation:

$$\frac{\partial C}{\partial t} = V_0 \vec{N}. \quad (6)$$

The evolution process of the latter is shown in the following formula:

$$\frac{\partial C}{\partial t} = ak \vec{N}. \quad (7)$$

In equation (7), a represents a positive constant, and k can accurately describe the bending degree of the curve. Then, the level set equation is constructed, and the level set function ψ is defined as hypersurface. Its expression is shown in the following equation:

$$\psi(x(t), t = 0) = \pm d. \quad (8)$$

In equation (8), $\pm d$ represents the symbolic distance between point x and surface C_0 . If x is located inside the surface, take a negative value; otherwise, take a positive value with derivative t , then equation (9) can be obtained.

$$\psi_t + \sum_{i=1}^N \psi_{X_i} X_{t_i} = 0. \quad (9)$$

When it is in the vertical state, there is the following equation:

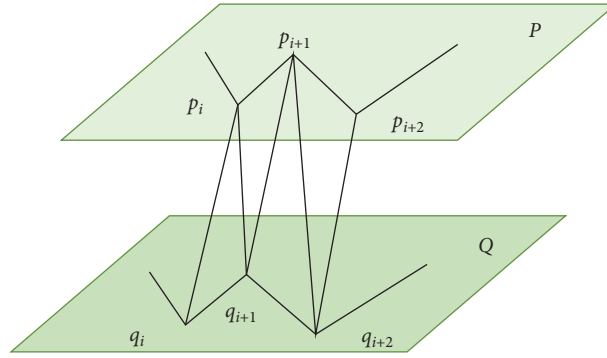


FIGURE 3: Structural diagram of basic triangle.

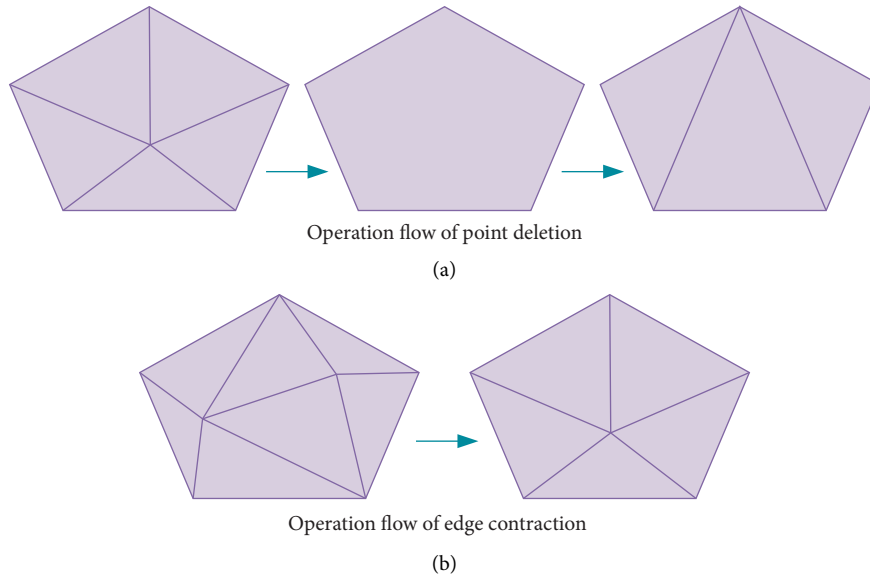


FIGURE 4: Schematic diagram of simplified processing operation of triangular mesh. (a) Operation flow of point deletion; (b) operation flow of edge contraction.

$$\sum_{i=1}^N \psi_{X_i} X_{t_i} = (\psi_{X_1} \cdots \psi_{X_N})(\psi_{1_t} \cdots \psi_{N_t}) = F(x(t))|\nabla\psi|. \quad (10)$$

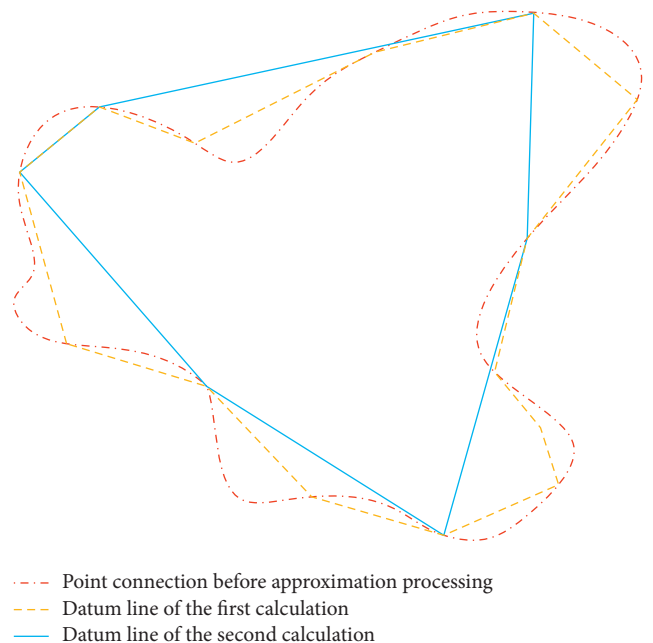
After evolution, the level set equation can be rewritten as in the following equation:

$$\psi_t + F|\nabla\psi| \approx 0. \quad (11)$$

In equation (11), ψ represents the level set function, then the contour curve of the target can be expressed as in the following equation:

$$\gamma(t) = (x \mid \psi(x, t) = 0). \quad (12)$$

3.2. 3D Reconstruction Technology Based on Contour Extraction and Sampling. Three-dimensional reconstruction technology is an important research direction of scientific visualization. This technology can process the two-dimensional pictures obtained by medical imaging equipment and then obtain the three-dimensional model of corresponding organs and tissues. Three-dimensional reconstruction technology mainly includes three different types, namely,



--- Point connection before approximation processing
 - - - Datum line of the first calculation
 — Datum line of the second calculation

FIGURE 5: Geometric principle of DP algorithm.

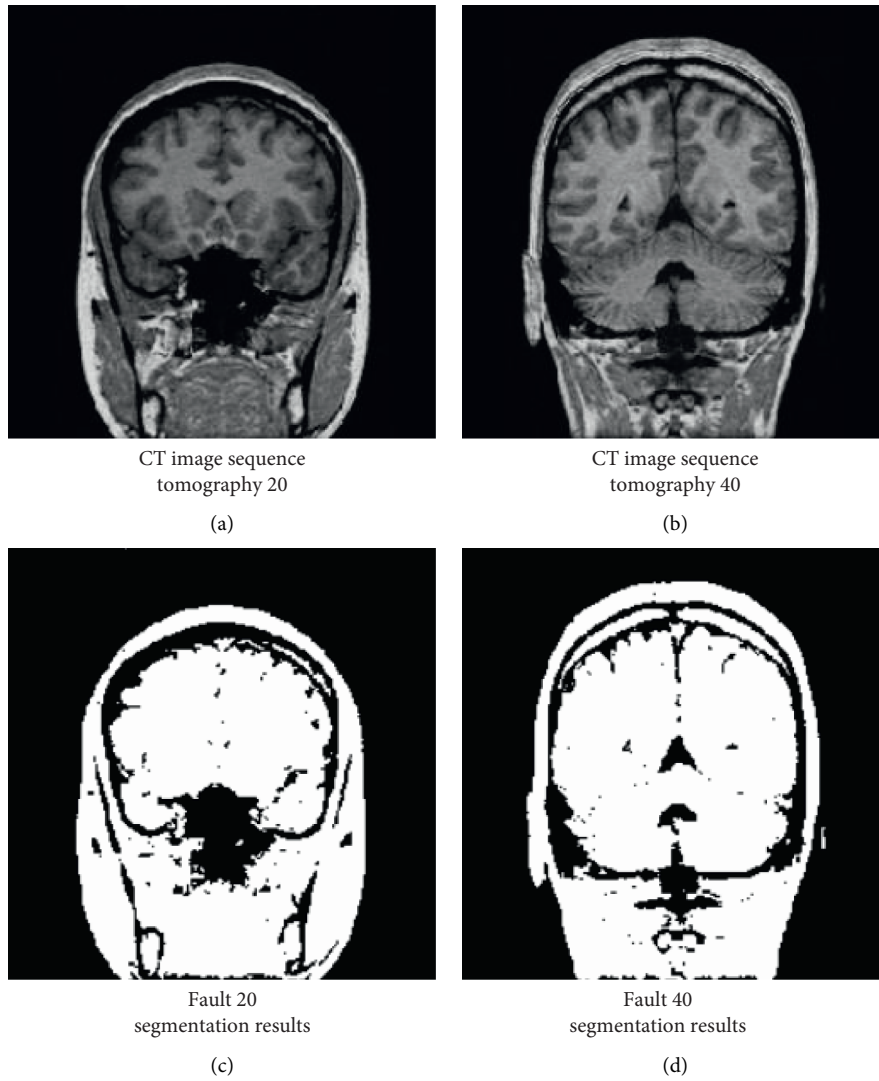


FIGURE 6: Extraction and segmentation results of head CT under different fault sequences. (a) CT image sequence tomography 20; (b) CT image sequence tomography 40; (c) fault 20 segmentation results; (d) fault 40 segmentation results.

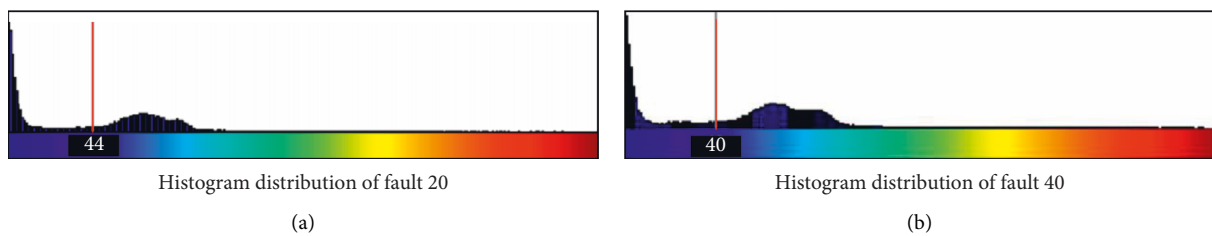


FIGURE 7: Histogram distribution of head CT under different fault sequences. (a) Histogram distribution of fault 20; (b) histogram distribution of fault 40.

surface rendering, volume rendering, and digital geometric processing, among which surface rendering has the best application effect. There is a triangulation method based on surface contour in surface rendering, which mainly uses the small plane with the shape of triangle or polygon to place it in the boundary contour line as a filler to form a complete and smooth object surface. Firstly, two feature point sets P

and Q are set, which are expressed as (13) and (14), respectively.

$$P = \{p_i \mid 0 \leq i \leq m\}. \quad (13)$$

In (13), the value of m is greater than or equal to 3, and $p_m = p_0$.

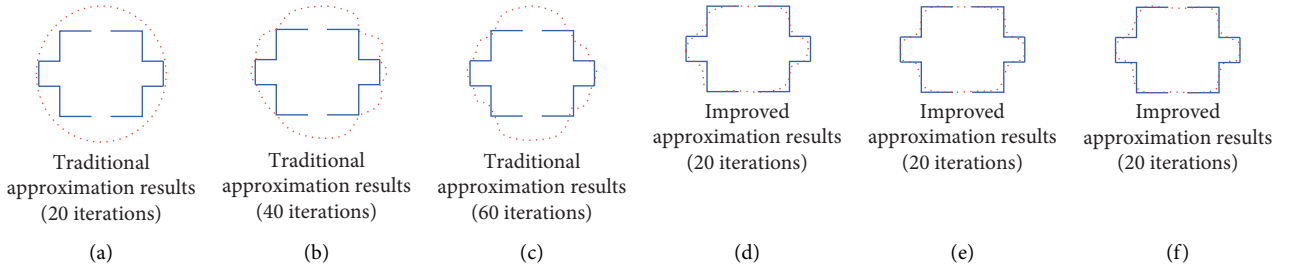


FIGURE 8: Comparison of approximation experimental results of two algorithms. (a) Traditional approximation results (20 iterations); (b) traditional approximation results (40 iterations); (c) traditional approximation results (60 iterations); (d) improved approximation results (20 iterations); (e) improved approximation results (20 iterations); (f) improved approximation results (20 iterations).

$$Q = \{q_i | 0 \leq j \leq n\}. \quad (14)$$

The value of n is also not less than 3, and $q_m = q_0$. P and Q are arranged in clockwise order in adjacent layers. Between the two, the three-dimensional reconstruction of the surface is completed by using triangulation. The applied triangle must meet four conditions at least one vertex P and one vertex Q in the triangle. If the intersection line between different triangles is not P_iQ_j such an edge, it must not intersect. Any intersecting edge P_iQ_j only belongs to two triangles that maintain adjacent relationship. Any point of a fault segment belongs to one of the adjacent triangles. A triangle meeting the four conditions can be called an acceptable surface as shown in Figure 3.

As can be seen from Figure 3, the two feature point sets are distributed up and down, and P and Q are the upper point set and the lower point set, respectively. Since the visible area of the triangle is small and contains a large number of triangular meshes, it needs to consume a lot of system resources every time rendering, so it is necessary to simplify the triangular mesh effectively. The simplified processing method mainly includes two types—point deletion and edge shrinkage. Its operation flow is shown in Figure 4.

Figure 4(a) shows the three basic steps of point deletion. First, select a point in the triangular mesh model to delete, then delete the edge with the target point as the vertex, and finally, all the holes are processed to obtain a new triangular mesh. Figure 4(b) shows the operation steps of edge contraction, which is mainly to move and merge the two endpoints located on the same edge to make it a point so that the original edge changes its connected endpoint and moves to the new vertex. In contrast, the operation of edge shrinkage is more convenient. It can complete the processing of two endpoints at one time and obtain a better simplification effect. Therefore, the application scope of edge shrinkage is far greater than that of point deletion.

There is a lot of data redundancy in the image and video data collected by widely deployed cameras. In video surveillance data, a large number of images and video data have time, space, and statistical redundancy. In order to reduce redundant data and shorten image processing time, contour sampling will be carried out using curve data compression algorithm (Douglas–Peucker, DP). The algorithm is a global sampling algorithm, which can obtain ideal and accurate sampling results based on vector or area offset.

The geometric principle of contour sampling by the DP algorithm is shown in Figure 5.

It can be seen from Figure 5 that during the sampling process of the DP algorithm, any arc of all curves in the polygon contour can realize the approximation with arbitrary accuracy, and the accuracy of the approximation can be described objectively and accurately according to the maximum distance between the point on the arc and the upper chord of the arc. The irregular polygon contour in Figure 5 is defined as V , and $V = \langle v_0, v_1 \dots v_n \rangle$. Let v_p represent any vertex in V , then its distance from any straight line $\overline{v_0v_n}$ can be expressed as $d(v_p, \overline{v_0v_n})$, and the corresponding approximation error is shown in the following equation:

$$\delta_{DP}(\overline{v_0v_n}, V) = \max(d(v_p, \overline{v_0v_n})) | v_p \in V. \quad (15)$$

In equation (15), δ_{DP} represents the approximation error, and then, a distance threshold ε used to represent the approximation accuracy is set. For all vertices in V , the distance between them and $\overline{v_0v_n}$ is calculated and compared one by one, and then, the vertex with the maximum distance is obtained, which is recorded as v_{p1} . If the corresponding maximum distance $d(v_{p1}, \overline{v_0v_n})$ is less than the established threshold, it can be determined that the irregular polygon V can be replaced by the straight line $\overline{v_0v_n}$, which marks the end of the decomposition. If not, it is necessary to decompose the subcontour recursively until the conditions are met.

4. Experimental Design and Analysis

4.1. Edge Extraction Results and Contour Model of CT Image Segmentation Technology. Compared with the traditional image segmentation technology, CT image segmentation technology based on human anatomy is more effective and can obtain excellent segmentation effect when the target object has a large contrast with the background. This method adopts the unitary global threshold segmentation technology to easily realize the effective extraction and segmentation of image edges in the process of pixel-by-pixel scanning and marking. See Figure 6 for details.

Figures 6(a) and 6(b) are the original CT images of the head when the fault sequence is 20 and 40, respectively. After using the CT image segmentation technology based on human anatomy, the fault segmentation results of Figures 6(c) and 6(d) are successfully obtained, which

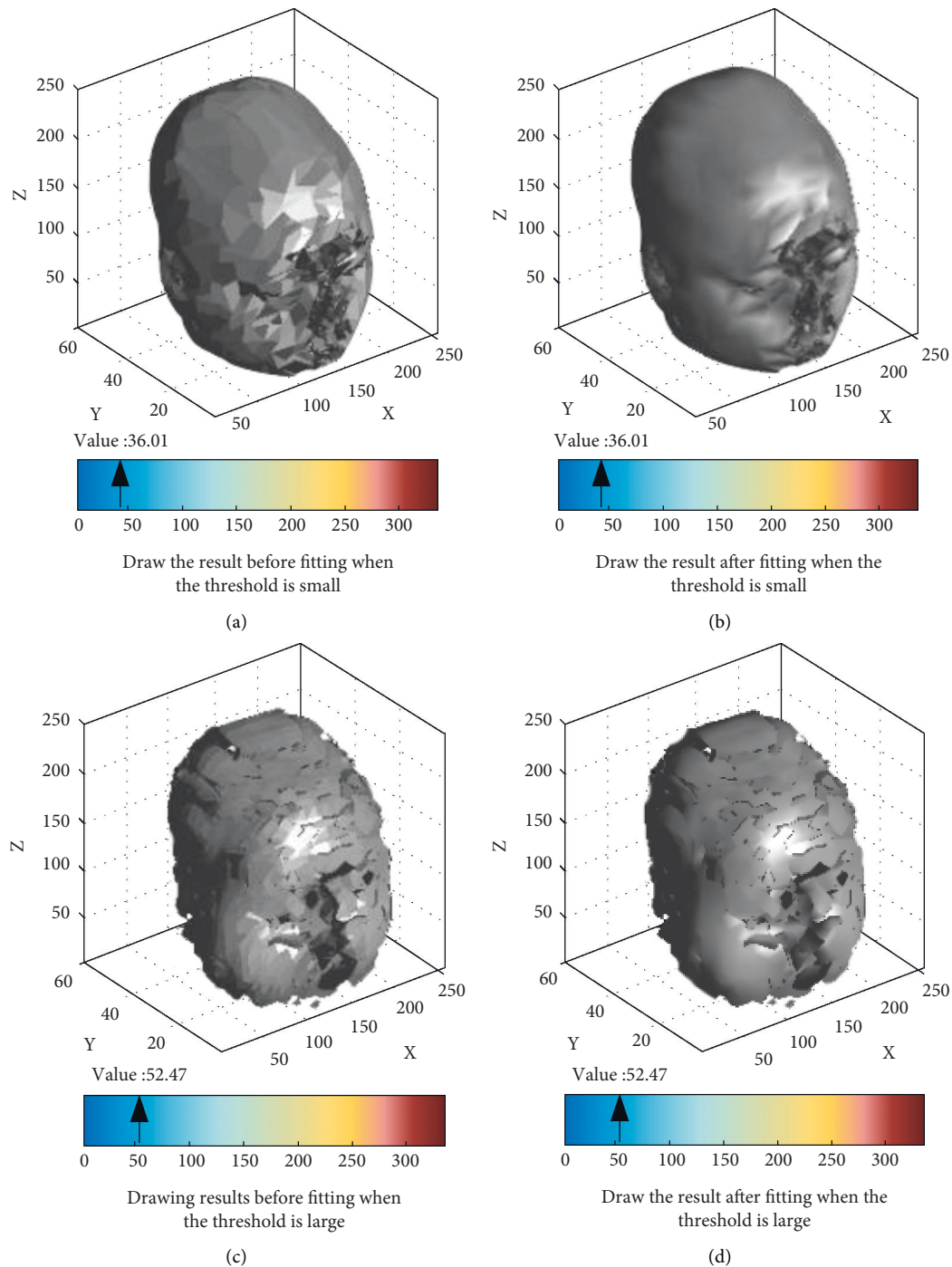


FIGURE 9: Comparison of surface fitting under different isosurfaces. (a) Draw the result before fitting when the threshold is small; (b) draw the result after fitting when the threshold is small; (c) draw the results before fitting when the threshold is large; (d) draw the result after fitting when the threshold is large.

correspond to two original images, respectively. It can be seen that the CT image segmentation technology based on human anatomy can obtain the more essential morphology and features of the target image and realize the significant improvement of image quality. In order to further explore the application effect of CT image segmentation technology, study and set different thresholds and analyze the gray histogram under different fault numbers, in order to obtain

the effectiveness of CT image segmentation technology in distinguishing the target image from the background. The results are shown in Figure 7.

According to Figure 7, under the influence of different thresholds, there are some differences in the distribution of threshold histogram, but the overall difference is small, which shows that CT image segmentation technology has significant universality and efficiency and can realize the

TABLE 1: Effects of different thresholds on 3D reconstruction time and effect.

Threshold	Reconstruction time (s)	3D reconstruction effect score
5	42.75	93.47
10	41.04	89.58
15	36.57	84.52
20	33.25	82.09
25	31.33	79.63
30	26.42	76.88
35	22.85	71.64
40	19.13	66.25
45	15.57	59.87
50	12.47	51.26

effective segmentation of target image and background without the interference of different thresholds. There is a close relationship between the extraction of the contour of the target image and the acquisition of the contour model. The higher the accuracy of the extraction of the former, the higher the quality of the latter. See Figure 8 for the convergence contour generated by different algorithms in approximation processing.

Figures 8(a)–8(c) are the approximation results of the traditional image segmentation technology. The difference lies in the number of iterations, which are 20, 40, and 60 in turn. The blue line represents the original contour of the target image, and the red line represents the approximate contour. According to Figures 8(a) to 8(c), with the increasing number of iterations, the similarity between the approximate contour and the original image shows a certain improvement, but there are still large differences between them. Figures 8(d)–8(f) show the approximation results of CT image segmentation technology based on human anatomy, and the iteration times are also 20, 40, and 60, respectively. By comparing and analyzing Figures 8(d)–8(f), it can be seen that there is no significant difference in the convergence contour generated under different iteration times, which is almost consistent, and the convergence contour is highly consistent with the original contour.

4.2. Surface Rendering and 3D Display Results. Set the layer spacing and point spacing between data to 3 mm and 0.418 mm, respectively, so that the data scale is $512 \times 512 \times 69$. Subsequently, the experimental results of surface rendering are obtained as shown in Figure 9.

Figures 9(a) and 9(c) are the surface drawing results before fitting, and Figures 9(b) and 9(d) are the surface drawing results after fitting. Compared with the rendering results before and after fitting, the head CT scanning image after fitting is clearer and the reconstruction effect is better. Comparing the fitting effects under different threshold levels, it can be seen that when the threshold level is low, the 3D reconstruction effect is better. If the threshold is too large, a large number of scattered fragments will appear in the 3D reconstruction results. In order to further explore the impact of threshold on 3D reconstruction results, this study

conducted 3D reconstruction experiments with different thresholds, recorded the time required for each experiment, and scored the 3D reconstruction results. See Table 1 for the results.

According to Table 1, with the increase in threshold, the time required for 3D reconstruction gradually decreases, indicating that the reconstruction efficiency shows an obvious improvement trend. Because a large number of discrete fragments will be displayed in the reconstruction results, the CT image of the head cannot be completely presented. Therefore, with the improvement of threshold and efficiency, the score of 3D reconstruction effect continues to decline.

5. Conclusion

At present, various diseases occur frequently in the medical community, and people's health is endangered by various diseases, among which head diseases are the most dangerous. It is very important to make effective diagnosis and treatment. In order to improve the level of medical diagnosis, this study will explore the mechanism of head CT image segmentation and three-dimensional reconstruction technology based on human anatomy and adopt different morphological operation methods and other measures, such as image expansion, image corrosion, point deletion, and edge contraction, in order to optimize head CT image segmentation technology and three-dimensional reconstruction technology. The experimental results show that under the influence of different thresholds, there are some differences in the distribution of threshold histogram, but the overall difference is small, which shows that CT image segmentation technology can effectively segment the target image and background. As for the three-dimensional imaging results themselves, the threshold has a great influence. The threshold is negatively correlated with the three-dimensional imaging time and the score of three-dimensional reconstruction results. This shows that in order to obtain better 3D reconstruction effect, it is necessary to ensure that the threshold is at a low level under the condition of reasonable time. The application effect of CT image segmentation technology and three-dimensional imaging technology based on human anatomy is good, which can provide reliable support for medical diagnosis and treatment.

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

The study was supported by the National Natural Science Foundation of China (grants 81901128 to WZY and 81901980 to ZDL).

References

- [1] C. Ye, W. Wang, S. Zhang, and K. Wang, "Multi-depth fusion network for whole-heart CT image segmentation," *IEEE Access*, vol. 7, Article ID 23421, 2019.
- [2] H. Xia, M. Ma, H. Li, and S. Song, "MC-Net' multi-scale context-attention network for medical CT image segmentation," *Applied Intelligence*, vol. 52, no. 1, pp. 1–12, 2021.
- [3] B. Biswas, S. K. Ghosh, and A. Ghosh, "A novel CT image segmentation algorithm using PCNN and Sobolev gradient methods in GPU frameworks," *Pattern Analysis & Applications*, vol. 23, no. 2, pp. 837–854, 2020.
- [4] M. Röding, C. Fager, A. Olsson, C. Von Corswant, E. Olsson, and N. Lorén, "Three-dimensional reconstruction of porous polymer films from FIB-SEM nanotomography data using random forests," *Journal of Microscopy*, vol. 281, no. 1, pp. 76–86, 2021.
- [5] S. A. Mason, I. M. White, T. O'Shea et al., "Combined ultrasound and cone beam CT improves target segmentation for image guided radiation therapy in uterine cervix cancer," *International Journal of Radiation Oncology, Biology, Physics*, vol. 104, no. 3, pp. 685–693, 2019.
- [6] G. K. Mourya, D. Bhatia, and A. Handique, "Empirical greedy machine-based automatic liver segmentation in CT images," *IET Image Processing*, vol. 14, no. 14, pp. 3333–3340, 2020.
- [7] D. Selvaraj, A. Venkatesan, V. G. V. Mahesh, and A. N. Joseph Raj, "An integrated feature frame work for automated segmentation of COVID -19 infection from lung CT images," *International Journal of Imaging Systems and Technology*, vol. 31, no. 1, pp. 28–46, 2021.
- [8] Y. Zeng, P.-H. Tsui, W. Wu, Z. Zhou, and S. Wu, "Fetal ultrasound image segmentation for automatic head circumference biometry using deeply supervised attention-gated V-net," *Journal of Digital Imaging*, vol. 34, no. 1, pp. 134–148, 2021.
- [9] M. Eckl, L. Hoppen, G. R. Sarria et al., "Evaluation of a cycle-generative adversarial network-based cone-beam CT to synthetic CT conversion algorithm for adaptive radiation therapy," *Physica Medica*, vol. 80, pp. 308–316, 2020.
- [10] E. Tappeiner, S. Pröll, K. Fritscher, M. Welk, and R. Schubert, "Training of head and neck segmentation networks with shape prior on small datasets," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, no. 9, pp. 1417–1425, 2020.
- [11] T. Yamamoto, T. Hasegawa, H. Hongo, and N. Amizuka, "Three-dimensional reconstruction of the Golgi apparatus in osteoclasts by a combination of NADPase cytochemistry and serial section scanning electron microscopy," *Histochemistry and Cell Biology*, vol. 156, no. 5, pp. 503–508, 2021.
- [12] M. Yushan, Y. Alike, A. Keremu, A. Abulaiti, P. Ren, and A. Yusufu, "Precise resection of macrodactyly under assistance of three-dimensional reconstruction technology: a case report," *Journal of Foot & Ankle Surgery*, vol. 59, no. 1, pp. 125–127, 2020.
- [13] X.-L. Xie, "Three-dimensional reconstruction based on multi-view photometric stereo fusion technology in movies special effect," *Multimedia Tools and Applications*, vol. 79, no. 13-14, pp. 9565–9578, 2020.
- [14] H. Wang, C. Zhang, Y. Song, B. Pang, and G. Zhang, "Three-dimensional reconstruction based on visual SLAM of mobile robot in search and rescue disaster scenarios," *Robotica*, vol. 38, no. 2, pp. 350–373, 2020.
- [15] M. Ruiz-Rodriguez, V. I. Kober, V. N. Karnaukhov, and M. G. Mozerov, "Algorithm for three-dimensional reconstruction of nonrigid objects using a depth camera," *Journal of Communications Technology and Electronics*, vol. 65, no. 6, pp. 698–705, 2020.
- [16] G. Jani, A. Johnson, U. Parekh, T. Thompson, and A. Pandey, "Effective approaches to three-dimensional digital reconstruction of fragmented human skeletal remains using laser surface scanning," *Forensic Science International: Synergy*, vol. 2, no. 4, pp. 215–223, 2020.

Research Article

Speckle Noise Removal Model Based on Diffusion Equation and Convolutional Neural Network

Siwei Nao¹ and Yan Wang² 

¹Medical Technology School, Qiqihar Medical University, Qiqihar 161000, China

²Basic Medical Science School, Qiqihar Medical University, Qiqihar 161000, China

Correspondence should be addressed to Yan Wang; shxwy@qmu.edu.cn

Received 16 April 2022; Accepted 16 May 2022; Published 15 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Siwei Nao and Yan Wang. 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 image denoising model based on convolutional neural network (CNN) can achieve a good denoising effect. However, its robustness is poor, and it is not suitable for direct noise removal tasks. Differently, the image denoising method based on the diffusion equation is more stable and has theoretical guarantees. In order to give full play to the advantages of CNN and diffusion equation in image denoising, this paper proposes a speckle noise denoising model via a combination of the two tools. Firstly, based on the mathematical model of speckle noise, a class of neural network speckle noise removal model which mixes residual learning and structure learning is proposed using image decomposition theory. Then, in order to solve the hyperparameter problem that the model depends on noise variance, a noise variance estimation algorithm based on a nonlinear diffusion equation is proposed. Finally, a speckle noise denoising model based on diffusion equation and CNN is obtained. Numerical simulation experiments verify the accuracy of the variance estimation algorithm and also the denoising effect and practical application value of the proposed method.

1. Introduction

In recent years, imaging technology has continued to develop and has been widely used in aerospace, geological remote sensing, and digital medicine. In the field of digital medicine, coherent imaging systems play a crucial role, especially the ultrasound imaging Taxt [1]. The principle of ultrasound imaging technology is to use ultrasound to scan tissues and organs in the human body and then collect and process reflected echo signals to obtain ultrasound images. Compared with other medical imaging technologies such as CT and MRI, ultrasound imaging has the advantages of being harmless to the human body and can be realized in real time. So, it can be widely used in clinical practice. However, due to the coherent characteristics of the ultrasonic imaging system, the scattered echoes of the ultrasonic technique may interfere and generate speckle noise, which reduces the signal-to-noise ratio (SNR) of the image and seriously affects the subsequent image postprocessing works. Therefore, it is very necessary to study the technology of speckle noise removal in ultrasound images.

In practical clinical applications, the logarithmic compression technology is used to compress the original echo signal to a suitable range, and the contrast of the signal is adjusted to enhance the darker details [2]. Loupas et al. [3] proposed the mathematical model of ultrasound image noise as follows:

$$f = u + \sqrt{u}N, \quad (1)$$

where u is the real image, f is the noise image, and N is the Gaussian white noise with mean 0. This noise model is used in many ultrasound image denoising models. Loupas et al. [3] proposed the adaptive weighted median filtering, and Dutt and Greenleaf [2] proposed adaptive unsharp mask filtering based on local statistics per pixel. At the same time, the model based on wavelet analysis was also applied to the speckle denoise problem of compressed ultrasound images. Abraham and Kadah [4] proposed a speckle noise removal model combining wavelet shrinkage and TV regularization to protect the boundaries and details of the image. Barcelos and Vieira [5] used the boundary control function to detect

high-level noise near the noise point to improve the fidelity term and proposed an adaptive boundary control variational model. After that, Krissian et al. [6] proposed the following fidelity term of the speckle noise removal model for ultrasound images:

$$E(u) = \int_{\Omega} \frac{(f-u)^2}{u} dx. \quad (2)$$

In 2011, Jin and Yang [7] proposed the following minimization function:

$$\operatorname{argmin}_{u \in BV(\Omega)} \left\{ \int_{\Omega} |Du| dx + \lambda \int_{\Omega} \frac{(f-u)^2}{u} \right\} dx. \quad (3)$$

In this work, Jin and Yang also demonstrated the theoretical properties of the model in (3). Since then, there were also some corresponding improved models, such as Barcelos and Vieira [5] and Hacini and Djemal [8, 9]. In recent years, the ultrasonic image denoising model based on partial differential equations has been newly developed. Zhang et al. [12] proposed a speckle noise removal method based on nonlinear diffusion equation in 2009. In 2014, Bhateja et al. [10] proposed a weighted diffusion filtering method and its improved model. In 2018, Zhou et al. [11] proposed a variable exponential diffusion equation for speckle noise removal in ultrasound images, described as follows:

$$\frac{\partial u}{\partial t} = di \, v \left(\frac{\nabla u}{1 + (|\nabla u_{\rho}|/K)^{\beta(u)}} \right), (x, t) \in \Omega \times (0, T),$$

$$\frac{\partial u}{\partial \vec{n}} = 0, (x, t) \in \partial\Omega \times (0, T),$$

$$u(x, 0) = f(x), x \in \Omega.$$

(4)

In equation (4), $K, \rho > 0, u_{\rho} = G_{\rho} * u, \beta(u) = 2 - 2|u|^{\alpha}/M^{\alpha} + |u|^{\alpha}$. The main idea is to change the diffusion type of different image feature regions through variable exponential function.

Since the 1990s, there have been many research results based on supervised learning in machine learning in the field of image processing. The main idea is to extract features from existing data to train the corresponding machine learning model [13]. The key part of the algorithm based on machine learning is to select which features in the image to represent the image, and traditional machine learning algorithms generally extract features manually. However, due to the limited features extracted manually, its image processing effect also has certain limitations. Deep learning algorithms, which have been highly praised in recent years, try to let computers automatically learn optimal features from the current datasets. Among them, the most representative of deep learning is the convolutional neural network (CNN), which automatically finds features in images through multiple convolutional layers. In 1980, Fukushima [14] did some CNN-related research, and in 1995, Lo et al.

[15] applied CNN to image processing. In 1998, Lecun et al. [16] proposed LeNet for digit handwritten recognition after the initial success in the last century because the training process of CNN requires a lot of computation, but the computer performance was limited at that time. It was not until 2012 that AlexNet proposed by Krizhevsky et al. [17] won the ImageNet Challenge that year with a great advantage, and deep CNN once again attracted great attention. After that, deep learning developed rapidly, and many new models appeared, such as VGG [18] and ResNet [19]. Let x_{in}, x_{out} denote the input and output of a neuron in the model, respectively, w and b are the learned parameters, $h(\cdot)$ is nonlinear activation functions, and then x_{out} can be expressed as follows:

$$x_{out} = h(w^T x_{in} + b). \quad (5)$$

A neural network consists of several neurons stacked into multiple layers, and information is propagated between layers. When the propagation direction is only forward propagation without feedback connection, the neural network is also called feedforward neural network. The first layer of a feedforward network is called the input layer, the last layer is called the output layer, and the middle layer is called the hidden layer. When the neural network has a very large number of hidden layers, it is also called a deep neural network, such as Stacked Autoencoders (SAEs), Restricted Boltzmann Machines (RBMs), and CNN.

At the same time, there are many algorithms that apply neural networks to the field of image restoration. One of the representative research results is the CNN model for image denoising. Jain and Seung [20] proposed a CNN image denoising framework, which achieved good results in both nonblind denoising and blind denoising. Vincent et al. [21] proposed stack sparse denoising. Harmeling [22] proposed a multilayer perceptron-based denoising network structure. In 2017, Zhang et al. [23, 24] used deep convolutional neural network (DnCNN) for residual image learning and then obtained denoised images. The two key points of the DnNN network are residual learning [19] and batch normalization. In the training process of the DnCNN model, the mean square error between the predicted residual image and the real residual image is used as the loss function of the model. The symmetric convolutional autoencoder proposed by Mao et al. [25] is composed of multiple convolutional layers and deconvolutional layers, which learns from noisy image to restore clean image from one end to other end. RED-Net extracts the features of the image through the convolution layer and avoids the influence of noise, and the subsequent deconvolution layer reconstructs the image according to the obtained features. Different from DnCNN, the loss function in RED-Net is chosen as the mean squared error between the predicted image and the real image.

Although deep learning has achieved good experimental results in image restoration, the model is less robust and lacks a sound theoretical explanation. At present, many scholars try to use mathematical theory to theoretically explain the deep learning framework. Many studies have shown that there is a certain connection between deep

learning and partial differential equations [26]. For example ResNet can be interpreted as a specific differential equations [27].

This paper proposes a speckle noise removal model for ultrasound images based on diffusion equation and CNN. In Section 2, based on image decomposition theory, a hybrid CNN speckle noise removal model is proposed, which is divided into two parts: image noise estimation subnetwork and image structure estimation subnetwork. In Section 3, an estimation algorithm of speckle noise level in ultrasound images based on diffusion equation is proposed and used as a parameter of denoising model based on diffusion equation and CNN. Finally, in Section 4, the model proposed in this paper is numerically realized, and some experimental results are given. And the experimental results demonstrate the effectiveness of the model in this paper.

2. Image Speckle Denoising Model Based on Hybrid CNN

According to the mathematical model of speckle noise, the variance of speckle noise in ultrasound images depends on the gray value of the image. In the area with a large gray value of the image, the noise variance is large, and in the area with a small gray value of the image, the noise variance is small. From a local point of view, when the fluctuation of the gray value of the original image is relatively smooth, the noise level is relatively similar. According to image decomposition theory, images can generally be decomposed into two parts: cartoon part and texture part [28]. The cartoon part contains the main structure and the gradually changing part of the image, showing a large smooth area. So its gray value is relatively close, and then it can be seen from the noise model that the variance of the noise fluctuates less. The texture part contains the small-scale detail information and noise of the image. According to the noise model, this part of the noise fluctuates greatly.

Denoising algorithms based on convolutional neural networks have made good progress in additive denoising. Zhang [23] proposed the DnCNN model, the main idea of which is to learn the noise of the image and subtract the noise image from the learned noise to obtain a noise-free image. Shen et al. [25] proposed the self-encoding structure RED-Net in the literature [25], which gradually compresses the noisy image information through convolutional layers and filters the noise information while retaining the structural information for the subsequent image reconstruction process. Based on the image decomposition theory, this paper combines the above two types of network structures, uses DnCNN to estimate the texture part of the noisy image, and uses RED-Net to estimate the cartoon part of the noisy image. A hybrid CNN is proposed to build a speckle noise removal model. The input of the model is a noise image, and the output is a restored noise-free image. As shown in Figure 1, the model is divided into two subnetworks: an image noise estimation network and an image structure estimation network. The image noise estimation network is a fully

convolutional network. The image structure design network is a skip-connected convolutional autoencoder. Finally, the two-part network obtains the output predicted image through multiple convolutional layers. The advantage of this network is that, based on the image decomposition theory, CNN is used to simultaneously remove noise and extract image structural features and finally integrate the noise information and structural information to obtain the predicted image.

In Figure 1, Conv is the convolution layer, Deconv is the deconvolution layer, BN is the batch regularization, and ReLU is the linear rectification activation function. Next, let f be the noise image, $\hat{\mu}$ be the image noise estimate, and \hat{s} be the image structure estimate. For the image noise estimation subnetwork, let $\Psi_R(\Theta_R, \cdot)$ be the learned mapping, where Θ_R is the parameter to be learned, then the input of the network is the noise image f , and the output is the estimation of image noise $\hat{\mu}$, which satisfies

$$\hat{\mu} = \Psi_R(\Theta_R, f). \quad (6)$$

As shown in Table 1, the noise estimation subnetwork is a fully convolutional network with 17 layers of convolution, the size of the convolution kernel is 3×3 , the stride is 1, and appropriate zero padding is done to ensure that each layer outputs an image has the same length and width as the input image. The first 16 convolutional layers of the network all need the ReLU activation function. The 2–16 convolutional layers have batch regularization, the number of feature maps output by the middle 1–16 layers is 64, and finally, a new image is obtained that is the same size as the original image.

In the process of designing the CNN structure, the loss function is a crucial part, and its role can be analogous to the energy functional in the traditional method. Different loss functions should be designed for different image processing tasks, such as the Softmax function in the commonly used image classification network and L^1 and L^2 regression in the target detection network. For the image denoising problem, the commonly used loss function is the L^2 norm or L^1 norm between the restored image and the noisy image. In order to make the parameters in the two subnetworks fully trained, the loss function of the model also includes the direct constraints of the two subnetworks. The loss function used in the hybrid CNN model proposed in this paper is

$$L(\Theta) = \frac{1}{N} \sum_{i=1}^N \|\Psi_R(\Theta_R, f_i) - (f_i - u_i)\|^2 + \|\Psi_S(\Theta_S, f_i) - u_i\|^2 + \|\Psi(\Theta, f_i) - u_i\|^2, \quad (7)$$

where $(f_i, u_i)_{i=1}^N$ is the N noisy images and no-noise images in the training set, $\Psi(\Theta, \cdot)$ is the mapping learned by the entire network, and Θ represents all parameters in the network.

Figure 2 shows the processing of a noisy image by the image noise estimation subnetwork in the hybrid CNN denoising model. The main purpose of the noise estimation subnetwork is to estimate the noise of the image. With the progressive number of network layers, the structural

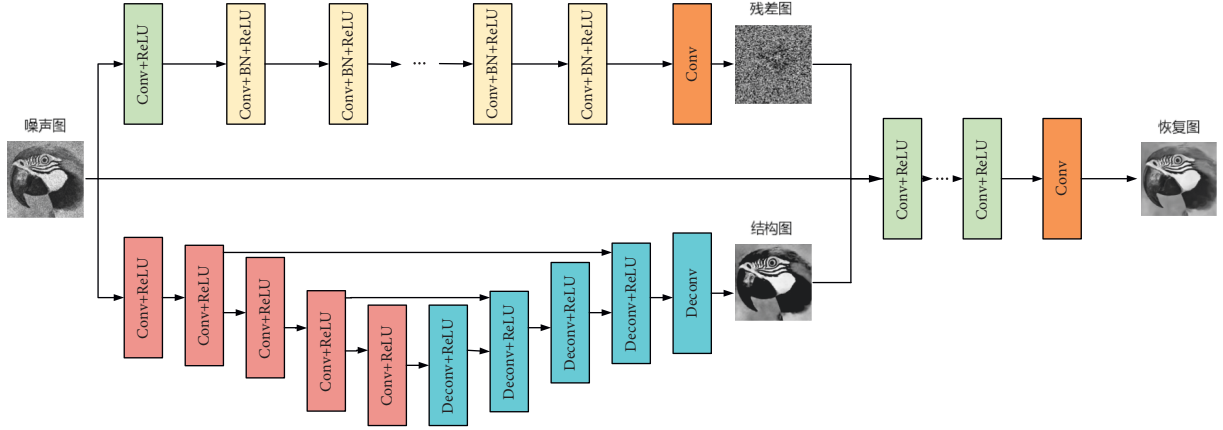


FIGURE 1: Schematic diagram of the hybrid CNN speckle noise removal model.

TABLE 1: Specific parameters of the network structure of the hybrid CNN speckle noise removal model.

Subnet type	Layers	Type	Number of input and output feature maps	Convolution kernel size and stride
Noise	1	Conv + ReLU	1 → 64	3 × 3, 1
	2–16	Conv + BN + ReLU	64 → 64	3 × 3, 1
	17	Conv	64 → 1	3 × 3, 1
Structure	1	Conv + ReLU	1 → 64	3 × 3, 2
	2–5	Conv + ReLU	64 → 64	3 × 3, 2
	6–9	Deconv + ReLU	64 (+64) → 64	3 × 3, 2
	10	Deconv	64 → 1	3 × 3, 2
Output	1	Conv + ReLU	1 + 1 → 64	3 × 3, 1
	2–6	Conv + ReLU	64 → 64	3 × 3, 1
	7	Conv	64 → 1	3 × 3, 1

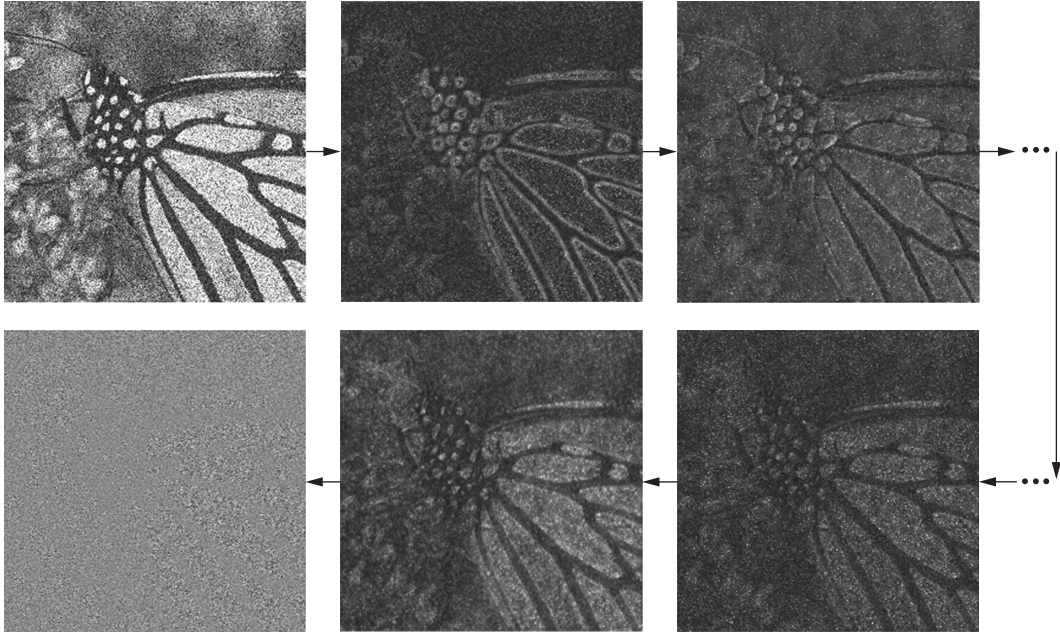


FIGURE 2: Processing of image noise estimation subnetwork.

information in the image becomes less and less, and the noise information becomes more and more. Figure 3 shows the processing process of the image structure estimation subnetwork. The image information is gradually compressed

by convolution downsampling. The structural information of the image is extracted, the noise information is removed, and the detailed information of the image is gradually restored in the process of upsampling.

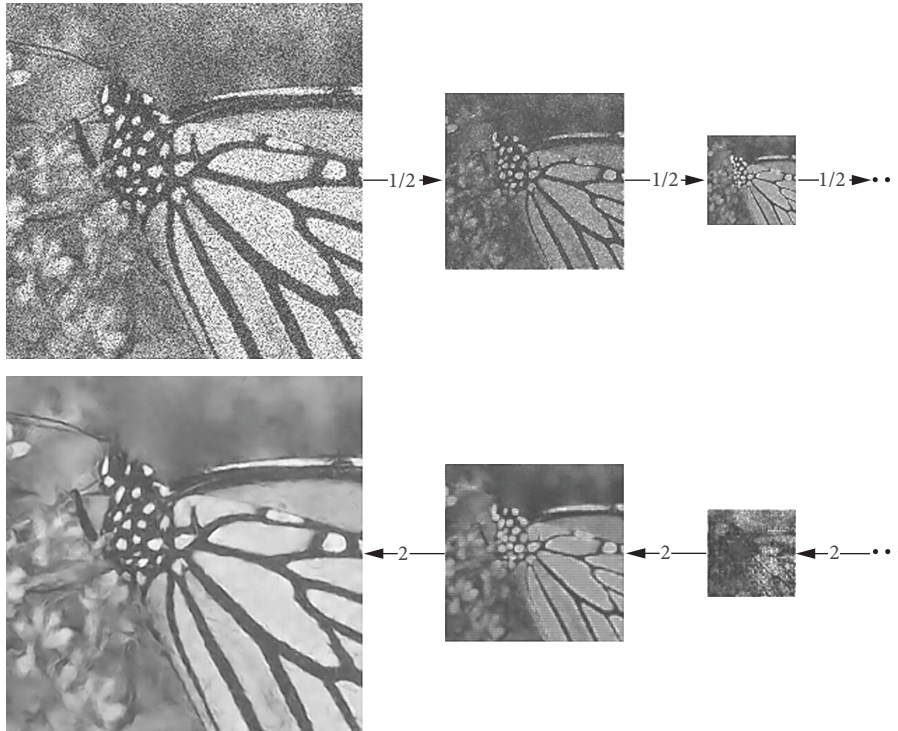


FIGURE 3: The processing of the image structure estimation subnetwork, $1/2$ and 2 are the multiples of image scaling.

3. Image Speckle Denoising Model Based on Diffusion Equation and CNN

CNN can achieve very good experimental results in image restoration, but it has a fatal disadvantage: the dependence on the noise variance parameter seriously limits its application in practical engineering problems. DnCNN and RED-Net need to be consistent in the image noise in the training process dataset, so these network structures can remove noise with known variance well, but they will not get a good denoising effect for images with unknown noise variance. Figure 4 shows a set of experimental results, trained on the training dataset with noise level $\sigma=25$, on DnCNN and tested on noise image with variances of 15, 25, and 35, respectively. From Figure 4, it can be found that, for the noise image with a noise variance of 15, the restoration result of DnCNN is too smooth, and the detailed information is lost. For the noise image with a noise variance of 35, the noise removal is insufficient, and there is still a lot of noise in the restoration result. For an image with a noise variance of 25, the denoising effect is very good; that is, the noise is removed, and the details are retained. In the actual application process, the variance of image noise is usually unknown, so directly using DnCNN to denoise may not necessarily get a good denoised image, which seriously limits the value of the neural network model in practical engineering applications. The same problem exists with the hybrid CNN model proposed in the previous section.

Aiming at the problem that the above CNN model relies on the noise variance, an intuitive outcome is to use noise images with different noise variances as the training dataset,

that is, mixed noise images. However, this has caused a series of problems; for example, the training process is more difficult, and the recovery effect of the final model is generally worse. Table 2 shows the PSNR results of the DnCNN model trained with different training sets. It can be found that, under different noise levels, the PSNR of the mixed case is lower than that of the single case.

Another method used in this paper to solve this problem is to predict the noise variance of noisy images in advance. The specific method is to use the diffusion equation-based ultrasonic image denoising model [11] to estimate the noise level. The model is not sensitive to noise variance due to the existence of Gauss convolution. Among the three key parameters (β, K, α) in the model, generally, $\beta = 1$, $K = 1$, and α varies between 1 and 2 with the noise level. But in fact, when the parameters α are selected within [1, 2], the impact on the optimal results of the model is far less than the impact of noise variance on neural network models. The specific nonlinear diffusion equation ultrasonic image speckle noise variance estimation method is given in the following.

Step 1. Diffusion equation preprocessing.

Here, f is the noise image, u is the noise-free image, and u_{pre} is the preprocessed image obtained by nonlinear diffusion equation. According to the work by Zhou et al. [29], u_{pre} can be handled accordingly.

Step 2. Noise variance estimation region selection.

Using the region detection operator in Guo et al. [12],

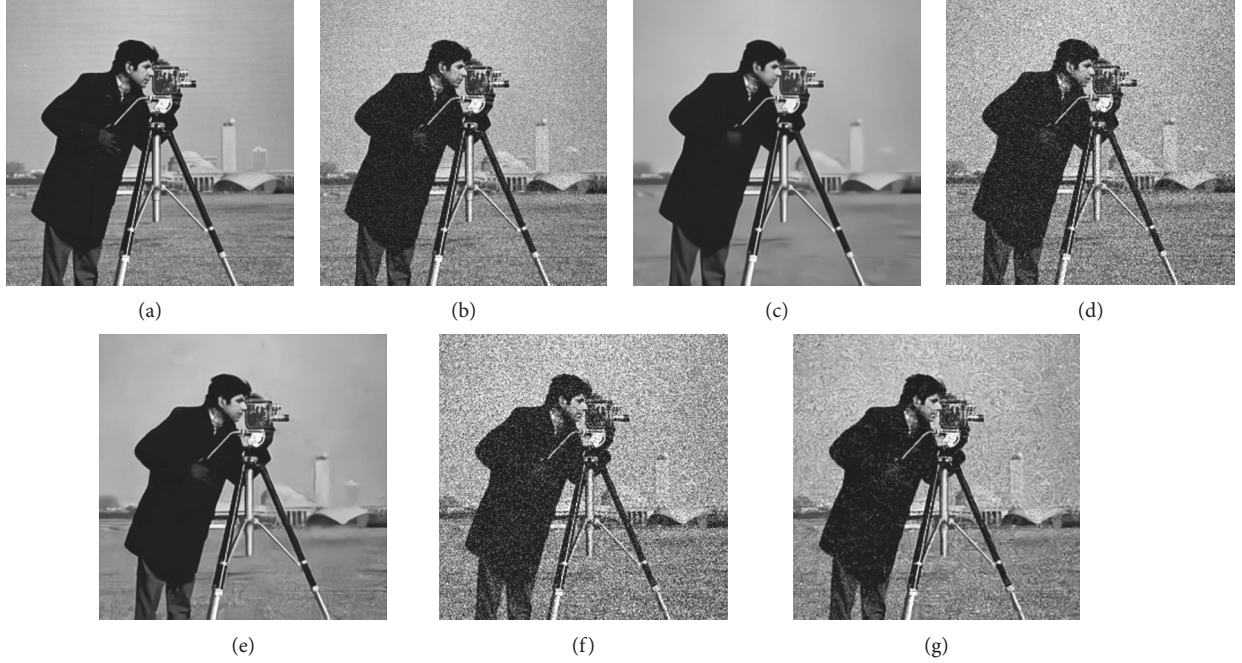


FIGURE 4: Restoration results of DnCNN on images with different noise variances. (a) The original image, (b) noise: $\sigma = 15$, (c) restored image, (d) noise: $\sigma = 25$, (e) restored image, (f) noise: $\sigma = 35$, and (g) restored image.

TABLE 2: PSNR results of DnCNN on different training sets.

Test image	I	II	III	IV	V	VI	VII	VIII	Average
Noise level $\sigma = 15$									
DnCNN (single) DnCNN (mix)	32.2732.01	34.64 34.22	32.62 32.11	33.37 33.12	31.72 31.54	33.29 32.93	35.02 34.83	33.39 33.01	33.29 32.97
Noise level $\sigma = 25$									
DnCNN (single) DnCNN (mix)	29.51 29.30	32.48 32.12	30.28 29.34	30.92 30.12	29.43 29.08	30.48 30.21	33.36 33.01	30.89 30.34	30.91 30.44

$$\beta(x) = 2 - \frac{2}{1 + |\nabla G_{\sigma} * u_{pre}|^2}. \quad (8)$$

The image structure u_{pre} in the preprocessing result $\tilde{\Omega}$ is detected. The image area is divided into two categories $\tilde{\Omega} = \{x \in \Omega, 0 \leq \beta(x) < 2\}$ and $\hat{\Omega} = \{x \in \Omega, 0 \leq \beta(x) < 2\}$.

Step 3. Noise variance estimation.

The sample variance of the noisy image f is calculated as

$$s_i^2 = \frac{1}{n_i - 1} \sum_{x \in \Omega_i} (f(x) - m_i)^2, \quad (9)$$

where n_i is the number of pixels in the subregion Ω_i and m_i is the mean of the noise image f in this region:

$$m_i = \frac{1}{n_i} \sum_{x \in \Omega_i} f(x). \quad (10)$$

To sum up, the specific process of the speckle noise denoising model for ultrasound images based on the

diffusion equation and CNN proposed in this paper is as follows: first, the noise image is preprocessed by the non-linear diffusion equation so as to estimate the noise variance. Denoising the image is treated as a parameter of the hybrid CNN model and taking the noisy image as input. It is worth mentioning that the whole process does not require any input parameters, thus improving the value of the model in practical applications.

4. Experimental Results and Analysis

This part presents the experimental results of the model in this paper and compares it with other models. The models compared with the model in this paper are the VA model [7], SRAD model [31], DnCNN, and RED-Net [25]. It is worth noting that the parameters of all models are adjusted to the optimal parameters according to the original paper, and the deep learning models are trained using the same training dataset. The optimization method used in the training process of the hybrid CNN model is ADAM [32]. The learning rate is 0.001, the number of batch training images is



FIGURE 5: Test image.

TABLE 3: PSNR results of different models on test images.

Test image	I	II	III	IV	V	VI	VII	VIII	Average
Noise level $\sigma = 2$									
Proposed	29.74	31.89	31.00	31.67	29.31	31.30	33.64	30.59	31.14
SARD	26.15	28.43	27.93	28.10	26.43	28.53	30.01	27.71	28.23
VA	26.32	28.96	28.01	28.31	26.77	28.69	30.28	27.82	28.51
DnCNN	29.65	31.88	30.95	31.66	29.23	31.24	33.63	30.57	31.10
RED-net	28.51	30.39	29.14	29.97	28.07	29.73	32.07	29.14	29.63
Noise level $\sigma = 3$									
Proposed	27.42	30.03	29.10	29.66	27.30	29.37	32.05	28.79	29.21
SARD	26.78	27.54	26.07	25.78	23.12	25.65	27.51	24.17	25.61
VA	24.11	25.74	25.48	25.98	23.77	25.87	27.98	24.76	25.91
DnCNN	27.36	29.85	29.09	29.64	27.14	29.30	31.94	28.77	29.14
RED-net	26.75	28.87	27.27	28.28	26.43	28.20	30.39	27.78	28.06
Noise level $\sigma = 4$									
Proposed	25.82	28.43	27.72	28.39	25.97	27.88	30.58	27.52	27.78
SARD	23.13	25.41	24.83	25.51	23.17	24.82	27.32	24.76	25.01
VA	23.51	25.62	25.14	25.86	23.67	25.21	27.88	25.03	25.62
DnCNN	25.65	28.33	27.78	28.12	25.81	27.62	30.40	27.48	27.64
RED-net	30.31	27.94	27.08	27.61	25.48	27.05	29.53	27.00	27.14

128, and the number of training times is 50. Figure 5 is the test image of the experiments in this section.

Table 3 gives the PSNR values of each model under different noise levels for 8 test images. In three sets of noise level experiments, the algorithm based on neural network is far superior to the traditional method in terms of PSNR index, which is 3 dB higher on average. It is worth mentioning that the PSNR value of the model in this paper is higher than other methods in most cases. Figures 6 and 7 show two sets of test result images. From the perspective of visual effects, although the traditional method can remove noise, the damage to the details of the image is very serious. The algorithm based on deep learning can not only achieve higher PSNR but also have better visual effects. The visual effects of the model in this paper are outstanding. Although

the model in this paper and the DnCNN model are very similar in terms of visual effects, there are certain differences in careful observation. For example, for the structure of the hat spike and mouth in Figures 6(b) and 6(f), it can be found that the image restored by the model in this paper is more detailed and closer to the original image. At the same time, a similar phenomenon can be found by comparing the petals in Figures 7(b) and 7(f). The results illustrate the advantages of our model in speckle noise removal.

Table 4 presents the experimental results of speckle noise variance estimation. Eight test pictures are tested, the noises $\sigma = 4, 9, 16$ are added to the test pictures, respectively, and then the noise level estimation method proposed in this paper is used to estimate the noise. It can be seen from Table 4 that, for different noise levels and different test

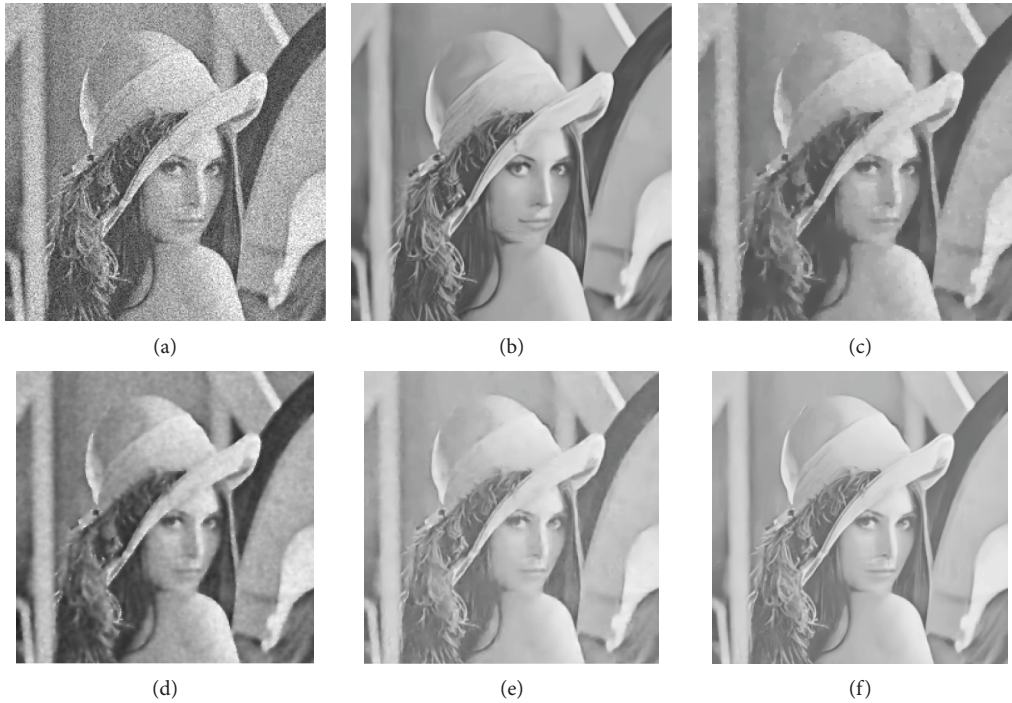


FIGURE 6: Results of test image II. (a) $\sigma = 2$. (b) The results of this model. (c) VA results. (d) SRAD results. (e) RED-Net results. (f) DnCNN results.

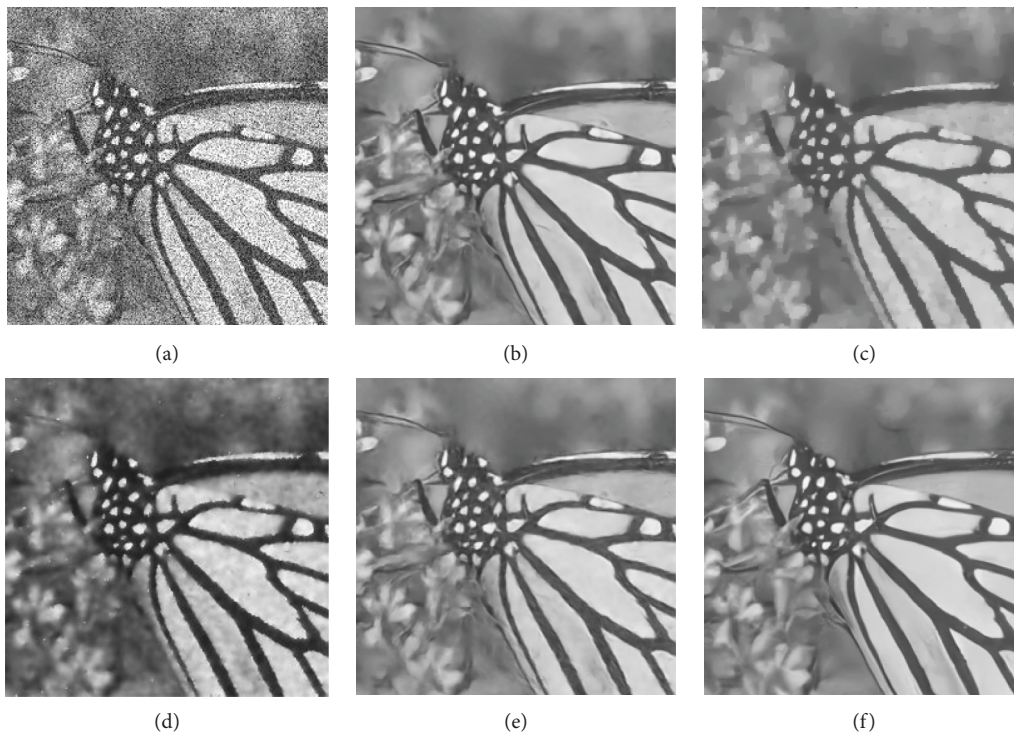


FIGURE 7: Results of test image VI. (a) $\sigma = 4$. (b) The results of this model. (c) VA results. (d) SRAD results. (e) RED-Net results. (f) DnCNN results.

images, the model proposed in this paper can predict the noise variance well, and the average error is about 0.06. This group of experiments verifies the effectiveness of the noise

level estimation method in this paper. Finally, the proposed model is applied to noise removal in real ultrasound images. Figure 8 presents a set of experimental results, from which it

TABLE 4: Noise variance estimation results on different test images.

Test image	I	II	III	IV	V	VI	VII	VIII	Average error
$\sigma = 4$	4.125	4.027	4.039	4.031	3.945	4.042	3.994	4.028	0.044
$\sigma = 9$	9.120	9.029	9.052	9.063	9.158	9.061	9.016	8.966	0.066
$\sigma = 16$	15.901	16.057	16.101	16.112	15.867	16.073	16.026	16.059	0.082

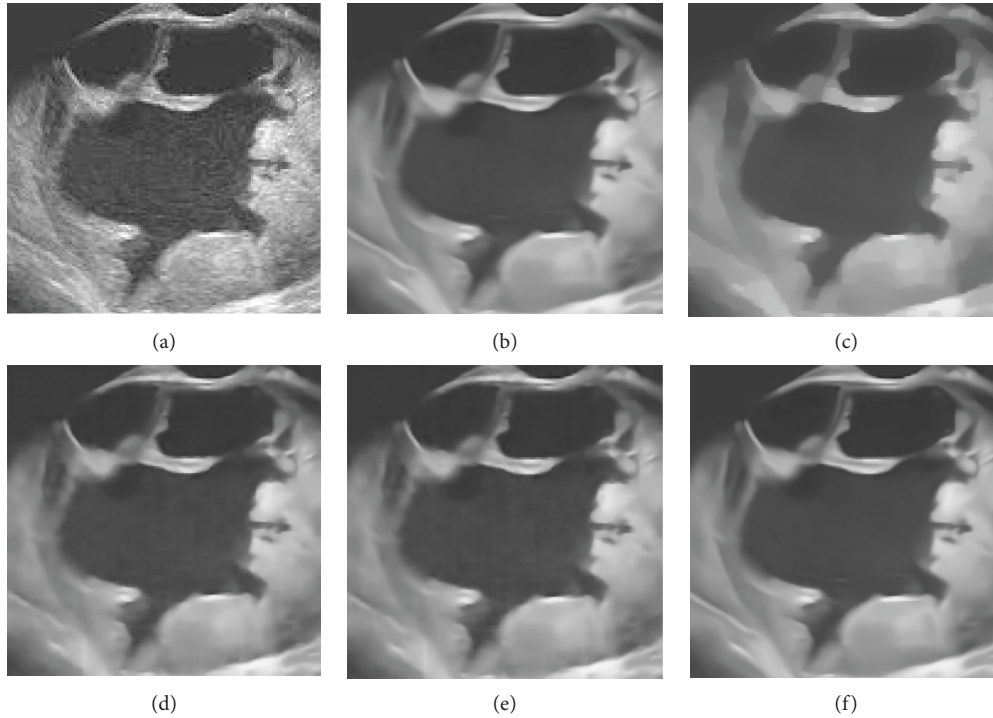


FIGURE 8: Results of real ovarian cancer ultrasound images. (a) The original image. (b) The results of this model. (c) VA results. (d) SRAD results. (e) RED-Net results. (f) DnCNN results.

can be seen that, compared with the traditional method and other neural network methods, the model in this paper can better restore the detailed information in the image and smooth the noise, which is better than other models in terms of visual effect. The model in this paper uses the diffusion equation method to estimate the noise level, thus avoiding the influence of the noise level parameter on the model, and in practice, the noise level has a great influence on the experimental results.

5. Conclusion

In this paper, an ultrasonic image denoising model based on diffusion equation and CNN is proposed, which fully utilizes the advantages of diffusion equation and CNN in image denoising. Based on image decomposition theory, the characteristics of texture and cartoon parts in ultrasound images are analyzed, and a hybrid CNN image denoising model is proposed, which is divided into noise estimation subnetwork and structure estimation subnetwork. Afterward, aiming at the problem that the convolutional neural network model depends heavily on the noise level, a method for estimating the noise variance of ultrasound images is proposed in this paper by taking advantage of the

insensitivity of the diffusion equation image denoising model to the noise level. The final experimental results show that the model in this paper can remove the noise in the image and restore the image structure information, and the model in this paper does not need any parameters and has great practical engineering application value.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the 2020 Qiqihar Science and Technology Research Planning Joint Guidance Project (no. LHYD-202020).

References

- [1] T. Taxt, "Radial homomorphic deconvolution of B-mode medical ultrasound images," in *Proceedings of the 12th IAPR*

- International Conference on Pattern Recognition (Cat. No.94CH3440-5)*, pp. 149–152, Jerusalem, Israel, October 1994.
- [2] V. Dutt and J. F. Greenleaf, “Adaptive speckle reduction filter for log-compressed B-scan images,” *IEEE Transactions on Medical Imaging*, vol. 15, no. 6, pp. 802–813, 1996.
 - [3] T. Loupas, W. N. McDicken, and P. L. Allan, “An adaptive weighted median filter for speckle suppression in medical ultrasonic images,” *IEEE Transactions on Circuits and Systems*, vol. 36, no. 1, pp. 129–135, 1989.
 - [4] B. A. Abraham and Y. Kadah, “Speckle noise reduction method combining total variation and wavelet shrinkage for clinical ultra-sound imaging,” *Biomedical Engineering*, pp. 80–83, 2011.
 - [5] C. A. Z. Barcelos and L. E. S. R. Vieira, “Ultrasound speckle noise reduction via an adaptive eedge-controlled variational method,” in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 145-151, Diego, CA, USA, October 2014.
 - [6] K. Krissian, R. Kikinis, C.-F. Westin, and K. Vosburgh, “Speckle-constrained filtering of ultrasound images,” in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, pp. 547–552, Diego, CA, USA, June 2005.
 - [7] Z. Jin and X. Yang, “A variational model to remove the multiplicative noise in ultrasound images,” *Journal of Mathematical Imaging and Vision*, vol. 39, no. 1, pp. 62–74, 2011.
 - [8] M. Hacini, F. Hachouf, and K. Djemal, “A new speckle filtering method for ultrasound images based on a weighted multiplicative total variation,” *Signal Processing*, Elsevier/North Holland, vol. 103, , pp. 214–229, 2014.
 - [9] F. Zhang, Y. M. Yoo, L. M. Koh, and Y. Kim, “Nonlinear diffusion in laplacian pyramid domain for ultrasonic speckle reduction,” *IEEE Transactions on Medical Imaging*, vol. 26, no. 2, pp. 200–211, 2007.
 - [10] V. Bhateja, A. Srivastava, G. Singh, and J. Singh, “A modified speckle suppression algorithm for breast ultrasound images using directional filters,” in *Proceedings of the ICT and Critical Infrastructure: 48th Annual Convention of Computer Society of India- Vol II*, pp. 219–226, Vishakapatnam, India, October 2014.
 - [11] Z. Zhou, Z. Guo, D. Zhang, and B. Wu, “A nonlinear diffusion equation-based model for ultrasound speckle noise removal,” *Journal of Nonlinear Science*, vol. 28, no. 2, pp. 443–470, 2018a.
 - [12] Z. C. Zhichang Guo, J. B. Jiebao Sun, D. Z. Dazhi Zhang, and B. Y. Boying Wu, “Adaptive perona-malik model based on the variable exponent for image denoising,” *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 958–967, 2012.
 - [13] G. Litjens, T. Kooi, B. E. Bejnordi et al., “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, no. 9, pp. 60–88, 2017.
 - [14] K. Fukushima, “Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.
 - [15] S.-C. B. Lo, S.-L. A. Lou, J. S. Jyh-Shyan Lin, M. T. Freedman, M. V. Chien, and S. K. Mun, “Artificial convolution neural network techniques and applications for lung nodule detection,” *IEEE Transactions on Medical Imaging*, vol. 14, no. 4, pp. 711–718, 1995.
 - [16] 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.
 - [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Proceedings of the International Conference on Neural Information Processing Systems*, vol. 60, no. 6, pp. 84–90, Long beach, CA, USA, February 2017.
 - [18] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” pp. 1–12, 2014, <https://arxiv.org/abs/1409.1556>.
 - [19] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.
 - [20] V. Jain and H. S. Seung, “Natural image denoising with convolutional networks,” in *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 769–776, Guangzhou, China, December 2008.
 - [21] P. Vincent, H. Larochelle, L. Lajoie, Y. Bengio, and P. A. Manzagol, “Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion,” *Journal of Machine Learning Research*, vol. 11, no. 12, pp. 3371–3408, 2010.
 - [22] H. C. Burger, C. J. Schuler, and S. Harmeling, “Image denoising: can plain neural networks compete with BM3D?” in *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2392–2399, Providence, RI, USA, June 2012.
 - [23] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising,” *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
 - [24] S. Ioffe and C. Szegedy, “Batch normalization: accelerating deep network training by reducing internal covariate shift,” in *Proceedings of the International Conference on International Conference on Machine Learning*, vol. 448-456, Lille, France, July 2015.
 - [25] X. Mao, C. Shen, and Y.-B. Yang, “Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections,” in *Proceedings of the advances in neural information processing systems 2016*, pp. 2802–2810, Barcelona, Spain, December 2016.
 - [26] B. Dong, Q. Jiang, and Z. Shen, “Image restoration: wavelet frame shrinkage, nonlinear evolution PDEs, and beyond,” *Multiscale Modeling and Simulation*, vol. 15, no. 1, pp. 606–660, 2017.
 - [27] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Data-driven discovery of partial differential equations,” *Science Advances*, vol. 3, no. 4, pp. 1–6, 2017.
 - [28] S. Osher, A. Sole, and L. Vese, “Image decomposition, image restoration, and texture modeling using total variation minimization and the h/sup-l/norm,” in *Proceedings of the 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, vol. 1, IEEE, Piscataway, NJ, USA, September 2003.
 - [29] Z. Zhou, Z. Guo, and B. Wu, “A doubly degenerate diffusion equation in multiplicative noise removal models,” *Journal of*

Mathematical Analysis and Applications, vol. 458, no. 1, pp. 58–70, 2018b.

- [30] Z. Q. Wu, J. X. Yin, and C. P. Wang, “Elliptic and parabolic equations,” *Springer Proceedings in Mathematics and Statistics*, vol. 119, no. 1, pp. 101–114, 2006.
- [31] Y. J. Yu and S. T. Acton, “Speckle reducing anisotropic diffusion,” in *Proceedings of the Conference on Signals Systems and Computers*, vol. 11, no. 11, pp. 1260–1271, Pacific Grove, CA, USA, May 2002.
- [32] D. Kingma and J. Ba, “Adam: a method for stochastic optimization,” in *Proceedings of the International Conference on Learning Representations*, vol. 1-13, San Diego, CA, USA, May 2015.

Research Article

Mask R-CNN-Oriented Pottery Display and Identification System

Chuantao Wei 

Hubei Academy of Fine Arts, Wuhan, Hubei 430205, China

Correspondence should be addressed to Chuantao Wei; 20181886@hifa.edu.cn

Received 19 April 2022; Revised 23 May 2022; Accepted 26 May 2022; Published 13 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Chuantao Wei. 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 pottery identification methods are time consuming and costly. In order to cater to more pottery industry needs, we propose a Mask R-CNN-based pottery identification method to build an automatic pottery identification system. We first improve the loss function of Mask R-CNN by using generalized intersection over union loss function, through the pattern of migration learning, to compensate for the disadvantage of a small collective amount of pottery data. For different types of pottery, we use the mask algorithm to enhance the features of the outer contour of the pottery. In addition, we use the minimum external matrix algorithm to accurately extract the outer contour bit pose features of pottery to improve the model's accuracy in recognizing the outer contour of pottery. To meet the testing conditions of pottery, with the support of potters and archaeologists, we make our pottery data set according to pottery categories. The experimental results prove that our method performs best in the comprehensive recognition accuracy of pottery, with the recognition accuracy above 90%. The recognition accuracy is also the best in pottery color decoration and grain decoration, and the grain recognition accuracy stays above 87%, which is better than other pottery recognition methods.

1. Introduction

Pottery is an important point in the history of human civilization. As early as the Neolithic era, primitive ancestors had mastered the art of kneading and firing pottery from clay. With the development of human civilization, ancient people continued to research and advance pottery technology, people experimented with the selection of clay types and the mixing ratio with water, and the shape of pottery evolved from practical cooking utensils to works of art. According to archaeologists, for the primitive ancestors, the production and use of pottery were mainly around the development of household items, and in later historical evolution the pottery that was buried in the soil for a long time still maintain the original appearance. The biggest advantage of pottery is that it can be used for a long time and is not easily graded in clay. So much so that pottery was mass produced as a craft in ancient times. After a long history, it has been preserved intact. What has been preserved in pottery is not only the skillful pottery of the generation of craftsmen but also the history and culture of the dynasty. The pottery provides a valuable reference for later archaeologists and antique enthusiasts.

According to historical records, in the development of ancient pottery, the ancients had an in-depth study of the materials, proportions, fire, and pottery art. The study of pottery in ancient times was a knowledge covering analytical chemistry, literature, art and art, sculpture, handicraft, and so on [1, 2]. Ancient potters had a different understanding of pottery materials, and different dynasties used different mainstream pottery materials, which confirms the different traditional cultures of each dynasty. And the emergence of colored pottery better reflects the lifestyle of ancient people. In addition to the different understanding of carving techniques by the potter of each dynasty, the pottery of each era has a unique shape style. Then later in the emergence of color painting, pottery artists will be celebrity in fusing poetry, songs, and paintings into the pottery, for the pottery to give the soul of culture and art. With the change of dynasties, the development of pottery also showed different characteristics with the rise and fall of dynasties. Pottery reached its artistic pinnacle during the Tang Dynasty, with excellent categories and techniques. Today, archaeologists and pottery enthusiasts have studied ancient pottery to some extent. The different shapes of pottery, different patterns, different carving arts, and so on

can determine the culture and lifestyle of the dynasties behind the pottery [3].

From the appearance of the pottery for quality and age judgment is the first stage of pottery identification, a more professional way to identify pottery is to start with the pottery material analysis. In order to obtain more comprehensive information on pottery, archaeologists resort to chemical analysis of pottery materials or physical processing methods. This has significant advantages for the quality analysis of pottery shards, and researchers mostly use multivariate calibration method and metrological analysis in the pottery material analysis data. Such delicate research work is a must for archaeologists, and for the amateur pottery enthusiasts, these pottery identification works are too costly and of low applicability [4, 5]. To make pottery identification more humane and intelligent, some researchers began to try to use machine learning methods for feature learning and classification of pottery categories. Machine learning methods are extremely demanding for manual labeling of features, the labor cost is large, and the recognition accuracy has a great upside. Considering the shortcomings of machine learning methods, some researchers began to try deep learning methods, the fragmented pottery features, and then performed neural network modeling on the fragmented features and achieved the purpose of learning pottery features by updating and iterating the neural network method [6]. The neural network approach can greatly improve the accuracy and speed of pottery recognition. In the subsequent pottery research, to adapt to the feature recognition needs of different pottery, a large number of researchers began to build different neural networks and verify the efficiency of neural networks through a large number of experiments.

We refer to different neural network models, and finally, we propose a Mask R- based pottery recognition method to build a pottery automatic recognition system. We improve the loss function of Mask R- by using generalized intersection over union loss function, through the pattern of migration learning, to compensate for the disadvantage of a small collective amount of pottery data. For different types of pottery, we use the mask algorithm to enhance the features of the outer contour of the pottery. To meet the test conditions of pottery, we make our pottery data set according to pottery categories with the support of potters and archaeologists. Finally, we experimentally demonstrate the effectiveness of our method.

The rest of the paper is organized as follows. Section 2 presents the history and research results of pottery identification research. Section 3 details the related principles and implementation details of the Mask R- based pottery recognition network. Section 4 shows the experimental data sets and the analysis of the experimental results. Finally, Section 5 summarizes our research and reveals some further research work.

2. Related Work

The style and texture of pottery are one of the important sources of information about pottery, and the social customs and trade information of the time can be

determined based on the appearance and shape of the pottery. In China's Tang Dynasty, potters would carefully select different clays according to the needs of their crafts. Pottery fired from different clays varied greatly in appearance and texture, which is a source of information for archaeologists to identify pottery. The sculptural styles pursued by potters varied from period to period, and the art of hand relief carving took on different styles with the rise and fall of dynasties. In the opinion of decorative researchers, the relief style of pottery represents the artistic trends of the time, which is also a great source of information for pottery identification. By the 1870s, the pottery industry began to take shape, and pottery research was gradually increasing, with pottery analysis and research becoming a popular industry. The traditional pottery identification industry relied heavily on experienced pottery experts, potters, and archaeologists. They use visual perception and empirical judgments to identify types of pottery, techniques, and chronological information. Traditional methods of pottery identification are time consuming and costly. To cater to the needs of more pottery industries, researchers have tried to adopt artificial intelligence to identify pottery information automatically and gradually started to build a pottery feature database to complete the pottery inventory and classification tasks by matching the feature database with additional samples. Considering that pottery artifacts are affected by external environmental factors in the process of preservation, the problem of degradation of pottery relief patterns and its texture occurs. Some researchers try to use deep learning methods to ensure the robustness of the pottery identification system. Some researchers have tried to build a pottery textual feature database by 3D scanning technology and classify the feature database by manual labeling. Some researchers try to seek the relevant factors for pottery classification from different neural network combination strategies, to enhance the sensitivity of neural networks to pottery pattern features.

To reduce the cost of artifact classification and inventory work, researchers in the literature [7] proposed the idea of building a database of artifact 3D features, and they scanned all the pottery artifacts in inventory in 3D and manually labeled each 3D data. The establishment of the database provided a solid data foundation for the later deep learning model building. In archaeology, for pottery from early times, most of the pottery cannot be preserved intact to this day due to weather and natural environment, and more of the pottery is in the form of shards in our research. Considering the featured study of pottery shards, literature [8–11] proposed different local feature extraction methods, mainly around the shape and contour of pottery shards, and feature mapping with the data set to obtain pottery sources. Researchers in the literature [12–17] also established a separate pottery color database and material features database to compensate for the lack of color features and material features in the previous data set. In the literature [18–20], to distinguish the details of pottery decorations, researchers established a database of pottery pattern reliefs, which mainly focuses on different pottery decorations for information annotation.

Researchers in the literature [21] found in the study of pottery shards that the appearance profile of the shard can influence the expert's overall judgment of the surface model of the pottery. The curvature of the appearance of the sherd profile, axis detection parameters, symmetry, and bending strength are all factors that must be considered in sherd studies in pottery sherd matching work [22]. In the literature [18], an analogical approach was proposed in the study of pottery decoration, where the researcher correlated the classification of pottery with the color decoration and relief patterns of pottery to build a 2D pottery decoration corpus and mapped the pottery shard decoration features to the database as a way to match to the information of the original pottery. Since the 2D corpus is not effective in 3D relief decoration, researchers in the literature [19] tried to integrate computer vision techniques into the identification of pottery reliefs, which have multiple parallel patterns or irregular patterns in numerous archaeological finds, and each pattern requires an independent database to be associated with it. Therefore, researchers in the literature [20] proposed a convolutional neural network-based pottery decorative template matching method, where they first preprocessed the pottery shards, then laser scanned the shards to obtain a model of the shard carving patterns, then segmented the model with hyperbolic pattern features, and finally trained independent segmentation features with convolutional neural networks to obtain a pottery decorative template identification model.

There are many different types of pottery, from domestic pottery items to pottery craft artworks, each with a unique pattern of decoration. The design of specialized pottery identification models for different categories of pottery is the mainstream research philosophy of researchers today. To ensure a faster pottery classification at a later stage, each pottery is manually labeled at the pottery laser scanning stage, and the source of the labeled information behind the pottery is directly accessible when extracting the pottery 3D model. Researchers in the literature [23] were inspired by the pyramid histogram and used SVM models to classify the visual features of pottery scans. Since there is a lot of room for optimization of machine learning methods in terms of accuracy and speed. Researchers in the literature [24] proposed an AlexNet-based pottery identification model. Researchers in the literature [25] built a pottery identification unit in the VGG11 network inspired by Google neural network to achieve pottery feature identification. The researchers in [26] proposed a pottery identification method based on ResNet18 and verified the effectiveness of the method through experiments.

3. Method

3.1. Basic Network. We investigated many neural network algorithms and performed experimental validation, and finally, we chose Mask R-CNN as the network base. Mask R-CNN belongs to the instance segmentation algorithm, and in pottery category identification, Mask R-CNN can perform pottery feature acquisition from pixel level and also instance

segmentation from the pottery 3D model and take different colors to mask over the target features [27]. Mask R-CNN belongs to a two-stage algorithm, where the first stage scans the sample data to generate weight extraction candidate regions and the second stage outputs target class masks on a recurrent convolutional neural network branch. In the 3D space, the information contained in each target mask is extracted in a refined way for the whole tau, which can cover the recognition target accurately [28].

Mask R-CNN is obtained by optimizing based on R-CNN, which is a network to mine more feature information from the target, and the authors use a region generation network to build a feature pyramid. To prevent overfitting by stacking the network too much, Mask R-CNN borrowed the VGG network proposed by Google, and the authors proposed the ResNet network based on the VGG network with the adaptive improvement of Mask R-CNN, which has the biggest advantage of using ROIAlign instead of the region of interest pooling. The structure of the Mask R-CNN network is shown in Figure 1.

3.2. Region Proposal Network. The region proposal network (RPN) is the core of the Mask R-CNN network. Its network structure is shown in Figure 2. At the pixel level, RPN will traverse each pixel point, and each pixel point will generate K anchor frames of different sizes accordingly, and each anchor frame will correspond to a background label independently. When the intersection ratio between the target real frame and the predicted frame is greater than the predefined maximum threshold, the predicted frame will be considered as the foreground. When the intersection ratio is less than the predefined minimum threshold, it will be considered as the background. When the intersection ratio is between the maximum and minimum thresholds, the prediction frame is considered the no-target case. The final output of RPN has $4K$ dimensions of location information and $2K$ dimensions of final output feature information. R-CN

Each target prediction box generated by the RPN network will exist in the pattern of Figure 3, where the red box represents the initial prediction box, the green box represents the real box, and the blue box represents the final prediction box.

If the pixel position of each anchor box needs to be determined, it needs to be calculated using the border regression algorithm, and the area between the prediction box and the real box obtained by the algorithm is the prediction generation area of the target. Each box has corresponding parameter coordinates $[x, y, w, h]$, where x and y denote the pixel coordinates of the top left corner of the anchor box, w denotes the width of the anchor box, and h denotes the height of the anchor box. After the initial prediction box is feature mapped, the final prediction box then has the following mathematical expression.

$$f(P_x, P_y, P_w, P_h) = (\widehat{G}_x, \widehat{G}_y, \widehat{G}_w, \widehat{G}_h) \approx (G_x, G_y, G_w, G_h). \quad (1)$$

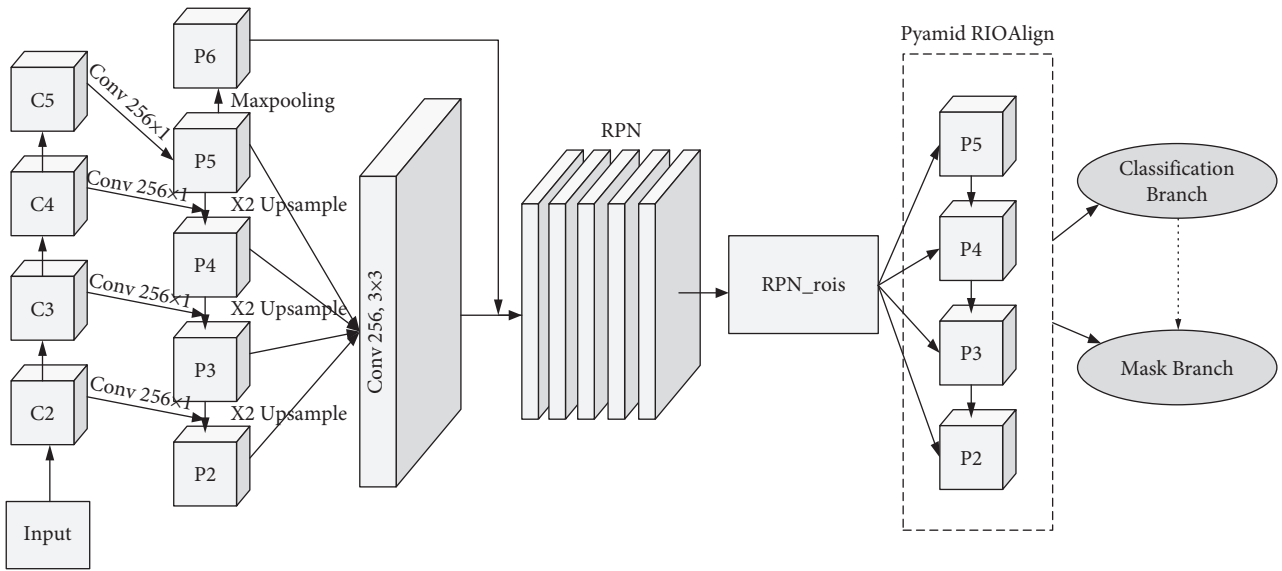


FIGURE 1: Mask R-CNN network.

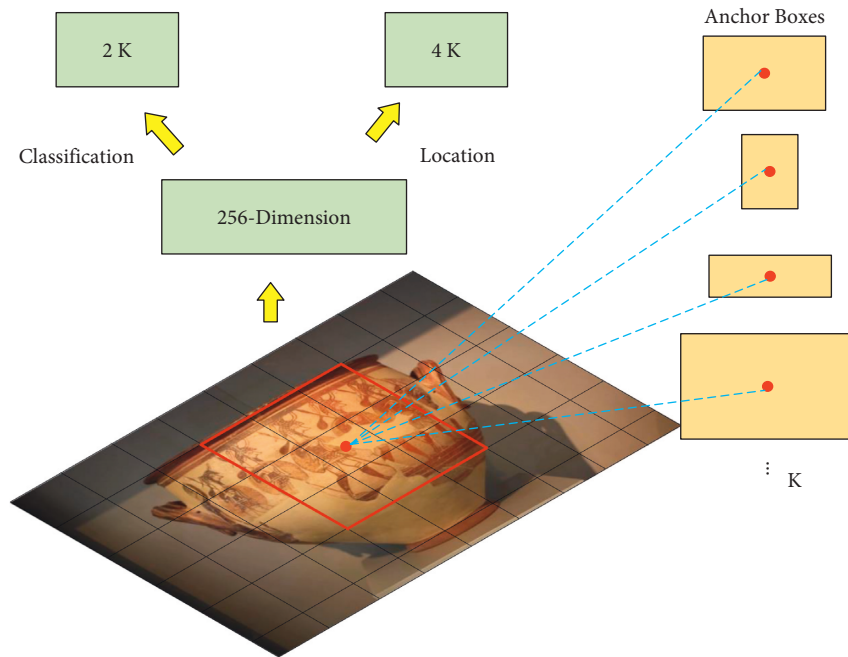


FIGURE 2: The principal structure of the region proposal network.

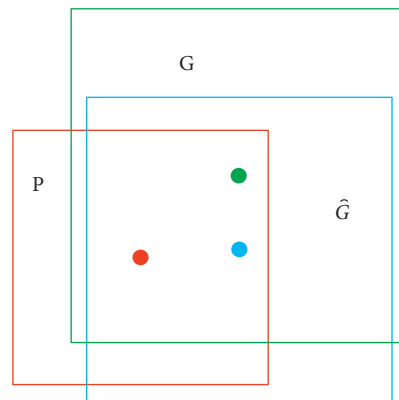


FIGURE 3: Region generation principle.

In the feature extraction process of the RPN network, the prediction frame will be automatically scaled according to the size of the real frame, and then the translation convolution will be calculated.

$$\begin{aligned}\widehat{G}_x &= P_w d_x(P) + P_x, \\ \widehat{G}_y &= P_h d_y(P) + P_y, \\ \widehat{G}_w &= P_w \exp(d_w(P)), \\ \widehat{G}_h &= P_h \exp(d_h(P)),\end{aligned}\quad (2)$$

where $d_*(P)$ denotes the initial predicted value and the true value is known, assuming $d_*(P) = w_*^T \Phi_5(P)$, $\Phi_5(P)$ denotes the feature vector of the input initial value, and w represents the parameter to be learned, and we also use the least-squares method to reduce the gap between the predicted value and the true value.

$$\text{Loss} = \sum_i^N (\mathbf{t}_*^i - \widehat{\mathbf{w}}_*^T \mathbf{f}_5(\mathbf{P}^i))^2. \quad (3)$$

W is obtained by bringing LOSS to a minimum value. The expression k of the ROI region can be obtained through the RPN network.

$$\begin{aligned}W &= \operatorname{argmin}_{w_*} \sum_i^N (\mathbf{t}_*^i - \widehat{\mathbf{w}}_*^T \mathbf{f}_5(\mathbf{P}^i))^2 + \lambda \widehat{w}_*^2, \\ k &= k_0 + \log_2 \left(\frac{\sqrt{wh}}{224} \right),\end{aligned}\quad (4)$$

where 224 denotes the pixel size of the target and k_0 denotes the layer where the ROI with area $w \times h = 224 \times 224$ is located. In addition, the value of k needs to be kept as an integer so that the large-scale ROI can extract large target features smoothly even on the low-resolution feature map, and the small-scale ROI can extract small target features from the high-resolution feature map.

3.3. Loss Function Optimization. The loss function of the Mask R-CNN network contains the classification loss function and regression loss function. The loss function of the Mask R-CNN model has an additional mask branch compared to the R-CNN network, and the mathematical expression is as follows:

$$L = L_{cls} + L_{box} + L_{mask}, \quad (5)$$

where L_{cls} denotes the classification loss function, L_{box} denotes the regression loss function, and L_{mask} denotes the mask loss function. Each independent loss function uses the L1 function and L2 function in predicting the edges and the true edges. Most of the loss functions take the cross-merge ratio as a constraint for calculating the overlap value between the prediction frame and the real frame, but the cross-merge ratio cannot reflect the form of overlap between the prediction frame and the real frame, resulting in the distance between the two cannot be accurately estimated, so we

propose the GIOU method, which is calculated as shown below.

$$\begin{aligned}IOU &= \frac{|A \cap B|}{|A \cup B|}, \\ GIOU &= IOU - \frac{|C/(A \cup B)|}{|C|}.\end{aligned}\quad (6)$$

Assume two arbitrary polygons A and B . C denotes the smallest closed object from the overlapping part of A and B . Generalized intersection over union (GIOU) is similar to the calculation of IOU with scale invariance. The value of IOU is always greater than the value of GIOU. When the prediction frame is far away from the real frame, the merge between them increases and GIOU decreases. The problem of not being able to judge the relative position between the predicted frame and the real frame can be solved by the way of this and the other.

3.4. Transfer Learning. Different deep neural networks have different training settings and different data set training durations. Our research focuses on the type identification of pottery, and pottery database construction is a huge project. The collective amount of pottery data that has been constructed at present is not large and the number of samples is limited. To verify the effectiveness of our method in a limited sample, we proposed a transfer learning method. We also refer to the migration model of feature parameters mentioned in the literature [29], which reduces the model training time and saves the cost of computer computing power. Migration learning contains the source domain and target domain in the mathematical expression, and their mathematical definitions are as follows:

$$\mathbf{D}(S) = \{\mathbf{x}, \mathbf{P}(\mathbf{x})\}, \mathbf{D}(t) = \{\mathbf{x}, \mathbf{P}(\mathbf{x})\}, \quad (7)$$

where $D(S)$ denotes the Source domain, $D(t)$ denotes the Target domain, x denotes the feature interval of the domain, and $P(x)$ represents the edge probability distribution corresponding to the feature space. We adopt the pottery data set as the auxiliary sample X_a and the pottery enhancement data set as the spatial sample X_b and assume that $Y = \{0, 1, 2, 3\}$, which contains 4 types of pottery as samples, the training data T and the test data S satisfy the following equations.

$$\begin{aligned}\mathbf{T} &\subseteq \{(\mathbf{X} = \mathbf{X}_b \cup \mathbf{X}_a) \mathbf{Y}\}, \\ \mathbf{S} &= \{(x_i^t)\}, \\ \mathbf{T}_a &= \{(x_i^a, c(x_i^a))\}, \\ \mathbf{T}_b &= \{(x_j^b, c(x_j^b))\},\end{aligned}\quad (8)$$

where the training set T is divided into two subsets T_a and T_b . $x_i^t \in \mathbf{X}_b$, where $i = 1, 2, \dots, k$, $x_i^a \in \mathbf{X}_a$, where $i = 1, 2, \dots, n$, $x_j^b \in \mathbf{X}_b$, where $j = 1, 2, \dots, m$. In the migration learning design of the pottery recognition network, we first normalize the pottery data set so that the data set satisfies a linear distribution. Then we update the sample weights of the pottery data set by using the public data set as network pretraining and comparing the pottery training results for error rate reset. For the correctly trained samples, we reduce

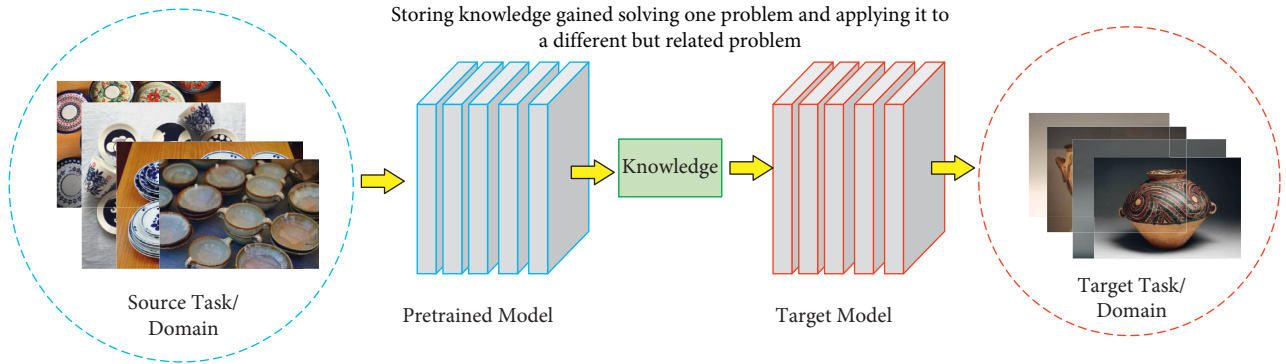


FIGURE 4: Transfer learning process.

the correlation weights, and for the incorrectly trained samples, we take the means of weight superposition. All pottery samples are updated iteratively until the error rate reaches within the specified range. In the pretraining process, we use ResNet as the base network to train the pottery for a phase of parameter calibration, adjust the parameters, and set the weights for the Mask R-CNN network according to the training results [30]. The detailed process of transfer learning is shown in Figure 4.

3.5. Pottery Identification System. Referring to a large number of pottery recognition studies, we choose to base on the Mask R-CNN network. Based on this, we introduced the GIOU loss function to enhance the extraction ability of the model on the generalized features of the target, and we chose Compute Unified Device Architecture to accelerate the model computation ability. To improve the pottery identification system, we added a laser scanner as the data source center of computer vision. The pottery scanning data samples are convolved to get the pyramid feature layer, and then the candidate regions are extracted on the feature layer, and the candidate regions are pooled and convolved to extract features. We introduce the GIOU loss function after the pooling layer and set 3 threshold criteria for the GIOU loss function and finally get the type features, grain location, and shape features of pottery by iterative training of pottery feature update and CUDA model acceleration. The model gets the pottery plane position after recognizing the pottery, but the laser scanner input is the three-dimensional information of the pottery, the pottery carving information requires spatial position features and angle features, and the bit pose estimation method mainly collects the pottery texture, edge shape, and other features. In the real environment, the imaging quality will be degraded due to factors such as lighting and texture, and the above geometric feature-based method is susceptible to great influence.

After the traditional deep learning target detection algorithm recognizes the pottery, it will be seriously affected by the lighting, texture, and complex background, etc., which leads to the difficulty of edge extraction and has a big impact on the pottery 3D feature prediction because there is no reinforcement of the edge information. Our method can also perform pixel-level feature segmentation for pottery in a

complex environment using the mask algorithm. The mask reinforces the pottery edge features and can better perform the task of locating the pottery 3D boundaries, and the edges of the already segmented image are extracted using the mask to generate pixel-level masks. We extract the mask obtained from the algorithm separately to generate the mask map and use the minimum external moment algorithm to accurately extract the pottery angle information, which can be used with the coordinate information obtained from the target recognition to obtain the pottery carving texture features [31]. The details of the pottery recognition system are shown in Figure 5.

4. Experiment

4.1. Data Set. There is no systematic pottery data set in the current pottery identification research.

To validate our pottery identification method, we made our pottery data set with the support of potters and archaeologists. To standardize the types of pottery, we set five major pottery categories at the early stage of pottery data set production, namely color pottery, faience, painted pottery, stamped pottery, and glazed pottery. Color pottery is mainly made of clay, ming, and material mixed with water in appropriate proportions, then shaped, dried, and fired. Faience is pottery made with iron oxide and manganese oxide as coloring elements. Painted pottery is pottery made by painting colors on the finished product after firing. Stamped pottery is pottery made by laying the clay on a bamboo mat or linen cloth and printing a pattern or cloth pattern. Glazed pottery is a compound of lead as an accelerant, after several firings of pottery, pottery external glaze is not only colorful but also not easy to fall off, commonly used in the language of crafts pottery production. According to the five major categories of pottery, we made different pottery data sets, respectively. The detailed data set information is shown in Table 1.

4.2. Experimental Settings. To ensure the independence and stability of the pottery identification system, we configured an independent host computer for the system, and all the integrated systems were developed with the host computer as the platform. In our experiments, we mainly configured the experimental environment with the Anaconda system. Considering the different requirements of the vision system

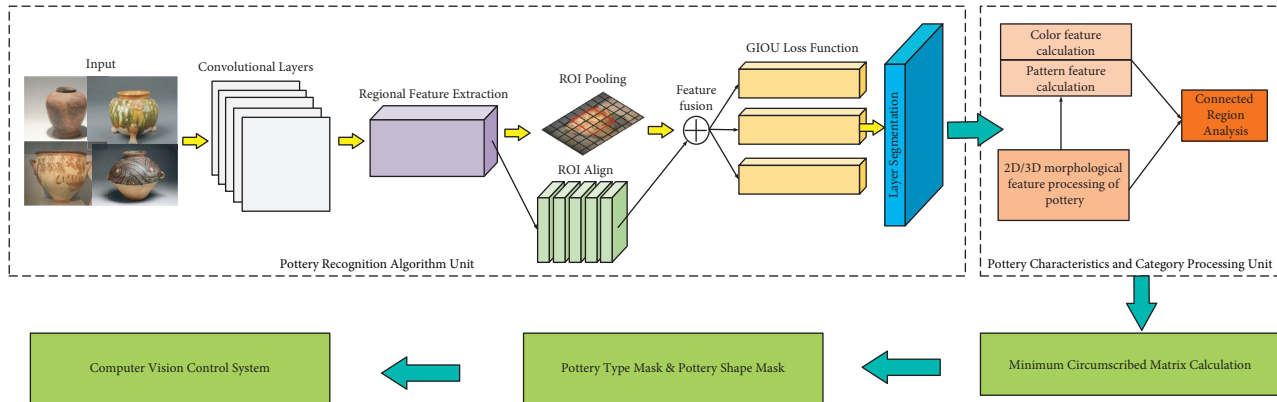


FIGURE 5: Pottery recognition system.

TABLE 1: Pottery data set classification and quantity.

	Train	Test	Total
Color pottery	15531	2011	17542
Faience	17323	1964	19287
Painted pottery	18036	2311	20347
Stamped pottery	16634	3002	19636
Glazed pottery	15321	1906	17227

and the pottery identification system for the programming environment, we configured multiple programming environments on the host computer to suit different needs. In the construction of the pottery identification neural network, we mainly use TensorFlow as the main framework. With the support of TensorFlow's powerful community module, our pottery identification network can be successfully built. The detailed training parameters are shown in Table 2.

For the method testing efficiency metrics, we chose recall (R) and precision (P) as general evaluation metrics. Where X_{TP} denotes correctly identified pottery, X_{FN} denotes unidentified pottery, and X_{FP} denotes misclassified pottery. The detailed mathematical calculation is shown below.

$$R = \frac{X_{TP}}{X_{TP} + X_{FN}},$$

$$P = \frac{X_{TP}}{X_{TP} + X_{FP}}.$$
(9)

4.3. Analysis of Experimental Results. To verify the effectiveness of our pottery recognition system, we compared machine learning methods and deep learning methods. Among the machine learning methods, we chose the SVM [32] algorithm, and among the deep learning algorithms, we chose the SSD algorithm [33] and the Faster R-CNN [34] algorithm. To ensure the independent validation relationship among the methods, we conducted five sets of experiments during the training process to independently validate the recognition efficiency of each group of methods for different classes of pottery. The detection results of each method are directly input into the statistical calculation part of the data set, and the final evaluation results are obtained by balancing

TABLE 2: Parameter settings.

Parameter	Value
Learning rate	0.01
Decay rate	0.0001
Momentum	0.9
Epoch	70
Iterations	1000
Dropout rate	0.5

the total number and quality of the data set. In the first phase of experiments, we validated all pottery data sets and compared the efficiency of our method with other methods, and the experimental results are shown in Table 3.

The X_{TP} in the experiment denotes the recognition efficiency per 500 pottery samples. From the experimental results in Table 3, it can be seen that the SVM method has the highest number of pottery misidentifications, accounting for one-fifth of the total. The recall rate is only 63.5% and the accuracy is not high. The pottery recognition efficiency is slightly better compared to SSD and Faster R-CNN methods, but there is still room for optimization. Our method has only 7 false identifications, and the pottery recognition accuracy reaches 98.3%. It is superior to other methods, which shows the superiority of our method in the first phase of experimental validation. In the second phase of the experiment, we verified the recognition accuracy of each class of pottery separately. Before the experiment started, we performed preprocessing operations on the data of no class of pottery to standardize the input format and sample frame rate of pottery to prevent the influence of data differences on the experimental results. The results of the second stage experiments are shown in Table 4.

From the experimental results of the second stage, it is clear that the recognition accuracy of traditional machine learning methods in various types of pottery is between 65% and 76%, which does not meet the practical needs in pottery classification work. Both SSD and Faster R-CNN methods keep the pottery recognition accuracy between 80% and 90%, and the pottery recognition efficiency is more stable. Our methods both maintain above 90% in pottery recognition accuracy, which proves the superiority of our methods in the second phase of experimental validation.

TABLE 3: Comparison of pottery identification experiments by different methods.

Method	X_{TP}	X_{FP}	X_{FN}	R (%)	P (%)
SVM	293	102	85	63.5	76.8
SSD	365	67	53	87.3	84.4
Faster R-CNN	383	35	29	90.9	91.6
Ours	463	7	5	97.1	98.3

TABLE 4: Comparison of recognition accuracy of different categories of pottery.

Method	Color pottery (%)	Faience (%)	Painted pottery (%)	Stamped pottery (%)	Glazed pottery (%)
SVM	70	73	76	69	65
SSD	82	85	86	80	79
Faster R-CNN	85	90	92	83	83
Ours	95	97	98	90	91

TABLE 5: Comparison of recognition accuracy of pottery decoration by different methods.

	Pattern decoration pottery (%)	Color decoration pottery (%)
SVM	65	72
SSD	72	78
Faster R-CNN	75	83
Ours	87	94

In the third phase of the experiment, to verify the recognition accuracy of the pottery identification system for three-dimensional pattern decoration and color decoration, we divided the pottery into two categories: pattern-decorated pottery and color-decorated pottery, where pattern-decorated pottery contains stamped and glazed pottery, and the remaining three categories are regarded as color-decorated pottery. The experimental results are shown in Table 5.

From the experimental results of the third stage, it is clear that all the methods have higher recognition accuracy in color pottery decoration than in pattern pottery decoration. Because patterned pottery decoration requires high 3D features, the traditional SVM method cannot capture pottery 3D features. SSD and Faster R-CNN outperform the SVM method in capturing 3D features, but still need to be optimized. Our method is the best performer in both color pottery decoration recognition and pattern pottery decoration recognition, and the pottery recognition accuracy is better than SSD and Faster R-CNN. It proves the effectiveness of our method in the third stage of experimental validation.

5. Conclusion

In this paper, we propose a Mask R-CNN-based pottery recognition method and build a pottery automatic recognition system. We first improve the loss function of Mask R-CNN by using generalized intersection over union loss function, through the pattern of migration learning, to compensate for the disadvantage of a small collective amount of pottery data. For different types of pottery, we use the mask algorithm to enhance the features of the outer contour of the pottery. In addition, we use the minimum external matrix algorithm to accurately extract the outer contour bit pose features of pottery to improve the model’s accuracy in recognizing the outer

contour of pottery. To meet the testing conditions of pottery, with the support of potters and archaeologists, we make our pottery data set according to pottery categories. The experimental results prove that our method performs best in the comprehensive recognition accuracy of pottery, with the recognition accuracy above 90%. The recognition accuracy is also the best in pottery color decoration and grain decoration, and the grain recognition accuracy is maintained above 87%. It proves that our method performs well in both feature capture at the two-dimensional level and three-dimensional level of pottery. Compared with traditional machine learning methods and deep learning methods, our method has higher recognition accuracy and better stability.

Compared with machine learning methods and deep learning methods, our method still has much room for improvement in recognition accuracy and recognition efficiency. Although it performs the best in the experiment of grain decorated pottery recognition. In future research, we will try to add a generative adversarial neural network as an auxiliary classification in the adversarial network to optimize the recognition of spatial pattern features during pottery pattern feature segmentation and improve the robustness and generalization of the network.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] T. Li, P. Li, H. Song, Z. Xie, W. Fan, and Q. Q. Lu, “Pottery production at the Miaodigou site in central China: archaeological and archaeometric evidence,” *Journal of Archaeological Science: Report*, vol. 41, Article ID 103301, 2022.
- [2] A. K. Marghussian, R. A. E. Coningham, and H. F. Nashli, “The development of pottery production, specialisation and standardisation in the Late Neolithic and Transitional Chalcolithic periods in the Central Plateau of Iran,” *Archaeological Research in Asia*, vol. 28, Article ID 100325, 2021.

- [3] C. Yuan, F. Wang, and S. Yuan, "Manufacturing techniques of sacrificial pottery from Jiaojia site, China, during the Dawenkou Culture," *Journal of Archaeological Science: Report*, vol. 40, Article ID 103238, 2021.
- [4] Y. Zhang, Y. Gao, J. Yang et al., "Patterns in pottery use reveal different adaptive strategies between lower and higher altitude regions on the Tibetan Plateau: chemical evidence from pottery residues," *Journal of Archaeological Science*, vol. 138, Article ID 105544, 2022.
- [5] B. Courel, J. Meadows, L. G. Carretero et al., "The use of early pottery by hunter-gatherers of the Eastern European forest-steppe," *Quaternary Science Reviews*, vol. 269, Article ID 107143, 2021.
- [6] J. Wang, W. Qian, and G. Chen, "Combining quantitative analysis with an elliptic Fourier descriptor: a study of pottery from the Gansu-Zhanqi site based on 3D scanning and computer technology," *Journal of Archaeological Science: Report*, vol. 36, Article ID 102897, 2021.
- [7] L. Gomes, O. Regina Pereira Bellon, and L. Silva, "3D reconstruction methods for digital preservation of cultural heritage: a survey," *Pattern Recognition Letters*, vol. 50, pp. 3–14, 2014.
- [8] M. Kampel and R. Sablatnig, "Rule based system for archaeological pottery classification," *Pattern Recognition Letters*, vol. 28, no. 6, pp. 740–747, 2007.
- [9] K. Son, E. B. Almeida, and D. B. Cooper, "Axially symmetric 3D pots configuration system using axis of symmetry and break curve," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 257–264, Portland, OR, USA, June 2013.
- [10] Z. Shunyi, H. Rongyong, W. Zheng et al., "Reassembling 3D thin fragments of unknown geometry in cultural heritage," *Photogrammetrie, Fernerkundung, GeoInformation*, pp. 215–230, 2015.
- [11] F. Zvietcovich, L. Navarro, J. Saldana, L. J. Castillo, and B. Castaneda, "A novel method for estimating the complete 3D shape of pottery with axial symmetry from single potsherds based on principal component analysis," *Digital applications in archaeology and cultural heritage*, vol. 3, no. 2, pp. 42–54, 2016.
- [12] M. Kampel and R. Sablatnig, "Color classification of archaeological fragments," in *Proceedings of the 15th International Conference on Pattern Recognition. ICPR-2000*, vol. 4, pp. 771–774, IEEE, Barcelona, Spain, September 2000.
- [13] F. Stanco and A. M. Gueli, "Computer graphics solutions for pottery colors specification[C]//Digital Photography IX," *SPIEL*, vol. 8660, pp. 194–203, 2013.
- [14] M. Farjas, J. G. Rejas, and T. Mostaza, "Deepening in the 3D modelling: multisource analysis of a polychrome ceramic vessel through the integration of thermal and hyperspectral information," in *Proceedings of the Computer Applications and Quantitative Methods in Archaeology (CAA)*, pp. 116–124, Beijing, China, April 2012.
- [15] P. Smith, D. Bepalov, and A. Shokoufandeh, "Classification of archaeological ceramic fragments using texture and color descriptors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, pp. 49–54, IEEE, San Francisco, CA, USA, June 2010.
- [16] M. Abadi, M. Khoudeir, and S. Marchand, "Gabor filter-based texture features to archaeological ceramic materials characterization," in *Proceedings of the International Conference on Image and Signal Processing*, pp. 333–342, Springer, 2012.
- [17] M. Makridis and P. Daras, "Automatic classification of archaeological pottery sherds," *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 5, no. 4, pp. 1–21, 2012.
- [18] G. C. Guarnera, F. Stanco, D. Tanasi, and G. Gallo, "Classification of decorative patterns in Kamares pottery," in *Proceedings of the SCCG 26th Spring Conference on Computer Graphics*, pp. 20–23, Slovak Republic, May 2010.
- [19] J. Zhou, H. Yu, K. Smith, C. Wilder, H. Yu, and S. Wang, "Identifying designs from incomplete, fragmented cultural heritage objects by curve-pattern matching," *Journal of Electronic Imaging*, vol. 26, no. 1, Article ID 011022, 2017.
- [20] J. Zhou, Y. Lu, K. Zheng, K. Smith, and C. Wilder, "Design identification of curve patterns on cultural heritage objects: combining template matching and CNN-based re-ranking," 2018, <https://arxiv.org/abs/1805.06862>.
- [21] A. Willis, X. Orriols, and D. B. Cooper, "Accurately estimating sherd 3D surface geometry with application to pot reconstruction," in *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop*, vol. 1, p. 5, IEEE, Madison, WI, USA, June 2003.
- [22] L. D. Angelo and P. D. Stefano, "Axis estimation of thin-walled axially symmetric solids," *Pattern Recognition Letters*, vol. 106, pp. 47–52, 2018.
- [23] T. Debrouille, S. Treuillet, A. Chetouani, M. Exbrayat, L. Martin, and S. Jesset, "Automatic classification of ceramic sherds with relief motifs," *Journal of Electronic Imaging*, vol. 26, no. 2, Article ID 023010, 2017.
- [24] A. Chetouani, S. Treuillet, M. Exbrayat, and S. Jesset, "Classification of engraved pottery sherds mixing deep-learning features by compact bilinear pooling," *Pattern Recognition Letters*, vol. 131, pp. 1–7, 2020.
- [25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <https://arxiv.org/abs/1409.1556>.
- [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*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [27] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, Venice, Italy, October 2017.
- [28] J. Zhao, D. Zhang, B. Shi, Y. Zhou, J. Chen, and R. Yao, "Multi-source collaborative enhanced for remote sensing images semantic segmentation," *Neurocomputing*, vol. 493, 2022.
- [29] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [30] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1717–1724, Columbus, OH, USA, June 2014.
- [31] M. I. Shamos, *Computational Geometry*, Yale University, Connecticut, CT, USA, 1978.
- [32] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 17, no. 1, pp. 113–126, 2004.
- [33] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, and C. Y. Fu, "Ssd: single shot multibox detector," in *Proceedings of the European conference on computer vision*, pp. 21–37, Springer, Amsterdam, The Netherlands, October, 2016.
- [34] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

Research Article

Protection and Planning of Historic Districts Based on Internet of Things Perception

Qing Zhao ^{1,2} and YunYing Ren²

¹College of Urban and Environmental Sciences, Northwest University, Xi'an 710000, China

²Architecture College, Xi'an University of Architecture and Technology, Xi'an 710000, China

Correspondence should be addressed to Qing Zhao; zq-xbxdx-1102@nwu.edu.cn

Received 17 February 2022; Revised 27 March 2022; Accepted 28 March 2022; Published 9 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Qing Zhao and YunYing Ren. 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 the Internet of Things era, the Internet of Things technology has been gradually accepted by people and gradually entered the category of block management, but the specific practical application scope is not wide enough. At present, the community management of residents in blocks generally adopts closed management, and there are some problems such as insufficient scope of property management, weak strength, and serious waste of human resources. Incomplete development of the hidden value is of cultural tourism blocks. At the same time, there are still some problems in the traffic management of blocks, such as incomplete monitoring scope and weak supervision. Therefore, this study proposes a bilevel programming model to solve the optimization scheme of block road network, adopts the complete information game model to analyze the imbalance between property and residents, and adopts the AHP fuzzy evaluation method to detect the current life happiness coefficient of urban residents and collect residents' suggestions for the future planning of historical blocks. Because the perception layer of the Internet of Things has the advantages of rapidity, accuracy, and unmanned operation, it can contribute a practical force to the protection and planning of historical blocks quickly and at low cost.

1. Introduction

In urban block planning, there are some problems in China, such as community closed pipe, crowded traffic routes, insufficient utilization of infrastructure, and so on [1]. However, in the era of big data information explosion in 2022, the digital information of urban blocks is bound to be on the right track, so the ancient survival mode of urban blocks has to lead to new changes. However, for some cities, some blocks have famous places of interest or rare regional characteristics, which is a pity to be destroyed. It is better to adopt Kevin Lynch's point of view that by protecting ancient buildings, we are preserving a sense of honor and history [2]. Therefore, the application of Internet of Things technology to the protection and planning of historical blocks has become the theme of urban planning in the new era.

The current situation of unreasonable infrastructure planning is the fatal factor of urban planning problems in

developing countries in the world [3]. The infrastructure system of the neighborhood has a long service life, and the impact on the environmental conditions lasts for a long time. Therefore, the strategic planning process of block facilities helps to realize the vision of environmental sustainability. Long-term investment in potentially destructive facilities may mean that sustainability priority principles are not included in planning and construction projects because these projects often follow the old planning model.

In the protection and planning of historical blocks, the first thing that should be built and improved is the community management mode. As early as 2000, China defined the community as the community of people's social life within a certain geographical scope for the first time in the article "Opinions of the Ministry of Civil Affairs on Promoting Urban Community Construction in China" [4]. Since then, the community has been formally determined as a unit of urban development. At present, our community is

basically closed, and the happiness coefficient of people's life is low, so we need to make reasonable planning and build the community into an open one. Starting from the perception layer of Internet of Things, this study explores the imbalance between property and residents by using the game model [5]. Second, AHP fuzzy evaluation is used to detect the current life happiness coefficient of urban residents and collect the focus of residents on community reform [6].

Second, we should study the rational development of cultural blocks on the premise of protection under the new technology. At present, China has tried to apply the Internet of Things technology to some historical and cultural blocks, but it is in an embarrassing situation that the application technology is not mature enough and the scope is not comprehensive enough. These blocks mainly take tourism industry as the development economy, and the activities of tourists from all over the world in this area have become one of the main sources of economic income for local cultural blocks. Therefore, in order to ensure the continuous input of sufficient passenger flow, we can design an intelligent bus system with the traffic demand of tourists as the guide [7]. At the same time, it is necessary to make reasonable planning for the allocation of cultural blocks in order to seek the result of maximizing the use of value [8].

Finally, with the technical support of the Internet of Things, the optimized road network of urban block traffic routes is designed. Simon, an outstanding scholar, put forward the concept of "satisfaction criterion," which defines that people are more inclined to make decisions that best meet their own needs, rather than theoretically optimal decisions [9]. Sun Yan designed a route selection model for passengers under the dual influence of rationality and preference [10]. Liu Kai designed a route selection model including three aspects: selection habits, traffic information, and time prediction [11]; Xu Hongli designed a path selection model for road risk avoidance and inertia decision [12]. Based on the achievements of the above scholars, this study puts forward a bilevel programming model to explore the optimization scheme of network route [13].

2. Protection and Planning of Residential Communities

Under the transformation of the Internet of Things perception technology, the future of residential communities is bound to be based on openness. The mobility of people in the community is strengthened, the freedom of residents' living activities is wider, and the work of people in all aspects of the property will be more flexible and meticulous. Community functions, including community population management, community business management, community sports management, community security management, and other management modules, will be improved accordingly, so as to seek the simultaneous improvement of residents' freedom of activities and happiness coefficient.

2.1. Research on Property Management under the Game Theory Model. After the opening of the block, the number of people

and vehicles in the community increased, which had an impact on the living comfort and the overall environment of the community, and the pressure on the property management department increased sharply. Therefore, it is necessary for the Internet of Things to fully coordinate the relationship among property management departments, community owners, and outsiders in the community. Among these three relationships, the property management department as the middle coordinator, its position is particularly important, so is the introduction of game theory knowledge to analyze the sustainable development of property management after the block opening.

In the analysis, we can know that both property enterprises and owners are rational economic people. Therefore, in community management, property managers and owners are principals and agents, respectively, so we can use the complete information game model to analyze the imbalance between owners and properties [14].

2.1.1. Model Assumptions. The research objects of the model are residents and property companies.

The strategy of the game model: assuming that the owner of the game party is $N1$, the strategy adopted by the owner is as follows: one is to pay the property fee regularly, which makes the property company gain profits and have the motivation to continue management, and the other is not to pay the property fee on time, so $N1 = (\text{pay the property fee on time, not pay the property fee on time})$. The property of the other parties participating in the game is $N2$, which can get property fees on time by actively providing services, but cannot collect property fees on time by passively providing services, that is, $N2 = (\text{high-quality services, low-quality services})$.

2.1.2. Modelling. Assuming that the income obtained by the property company under normal operation is M and the extra income obtained by the property fee collected by the property company is E , if the cost incurred by the property company in providing high-quality services for the owners is $C1$ and if the services provided by the property company are low cost, then the cost incurred by the property company at this time is $C2$, and at this time, $C1 > C2$.

When the property company provides high-quality services, such as good control of foreigners, timely provision of on-site warranty services, and property safety of owners, the community residents will generate income and assume $R1$. When the property company provides negative business services to residents, the income generated by residents is $R2$. Once residents refuse to pay property fees on the grounds of poor service of property companies, the amount will be used for residents' consumption of other matters, and the resulting income will be E . In terms of residents' income, the income of residents enjoying high-quality services must be greater than that of receiving low-quality services, that is, $R1 > R2$.

In this asymmetric nonzero game, no matter whether the property company provides quality management work or not, it will first consider delaying the payment of property

TABLE 1: Game between owners and property management companies.

		Residents, N1	
		Regular payment	Deferred payment
Property company, N2	Quality service	$M + E1 - C1, R1$	$M - C1, R1 + E1$
	Low-quality service	$M + E1 - C2, R2$	$M - C2, R2 + E2$

fees. At this time, the income of residents is R or $R2 + E2$. Property companies usually adopt the negative service mode, that is, $M + E1 - C2$, as given in Table 1.

2.1.3. Model Solving. Let α ($0 < \alpha < 1$) be the discount factor between residents and properties. If residents fail to pay fees on schedule, the benefits obtained by the owners are $R1 + E2$, not the income of the unit $R1$. The owner's deferred payment behavior will lead to the inability of the property to provide high-level property management work, and the owner can get the benefit of $R2 + E2$. Residents choose to pay on time according to the service quality of the property and should meet the following conditions:

$$(R1 + E2) + \alpha(R2 + E2) + \alpha^2(R2 + E2) + \dots \leq R1 + \alpha R1 + \alpha^2 R1 + \dots,$$

$$\text{Available after transformation } (R1 + E2) + \frac{\alpha(R2 + E2)}{1 - \alpha} \leq \frac{R1}{1 - \alpha}. \quad (1)$$

When $\alpha \geq E2/R1 - R2$ appears, it indicates that the community property provides high-quality services, so the owner will not take the initiative to delay payment.

As for that behavior of property company, if the property company chooses nonquality service when serving the owner, at this stage, the benefit that the property management company can get is $M + E1 - C2$, not $M + E1 - C1$. The behavior of the property management company will cause the owner to delay payment, which makes the income of the property management company $M - C2$ in each stage. If the owner can pay on time, the property management company will not adopt the strategy of low-quality service. It can be expressed as follows [15]:

$$(M + E1 - C2) + \alpha(M - C2) + \alpha^2(E - C) + \dots \leq (M + E1 - C1) + \alpha(M + E1 - C1) + \dots, \quad (2)$$

i.e.,

$$(M + E1 - C2) + \frac{(M - C2)\alpha}{1 - \alpha} \leq \frac{M + E2 - C1}{1 - \alpha}. \quad (3)$$

Solution is as follows:

$$\alpha \geq \frac{C1 - C2}{E1}. \quad (4)$$

When $\alpha \geq C1 - C2/E1$, residents pay property fees on time, and property companies provide high-quality services for residents.

According to the knowledge of game theory, if the number of times of the same result in the experiment reaches a certain value, then the payment vector can be obtained by a subgame equilibrium. In property management, if α can

satisfy $\alpha \geq E2/R1 - R2$ and $\alpha \geq C1 - C2/E1$, the trigger strategy is a subgame equilibrium of infinite repeated game, while the service management of property and the on-time payment of residents are the equilibrium results of each stage, in which α represents the patience or concern of participants in property management.

2.2. Residents' Satisfaction under AHP Fuzzy Evaluation. At present, community construction services are mainly concentrated in three aspects: social welfare services for vulnerable groups, collective routine services, and convenient services [16].

When the community becomes open, the flow of foreign personnel and vehicles continues to increase. The safety risk factors are gradually improved. At the same time, property management immediately changed from closed management service to open service. During this period, there will inevitably be inadaptability and neglect of work, so the corresponding service quality provided by the property to residents will also be impacted. Therefore, it is of great practical significance for the future planning of the block to design a scientific evaluation model based on the perception layer technology of the Internet of Things and explore the true inner thoughts of residents.

2.2.1. AHP Principle of Index Selection for the Evaluation Model

- (1) People-oriented principle: we should take the residents' own interests as the core of all works and fully meet the material and spiritual needs of residents [17].
- (2) The scientific nature of the evaluation index: it is the inevitable requirement of scientific and effective research results that we carry out rational screening on various needs of residents through comparative experiments and comparisons.
- (3) Systematicness of evaluation index: community management includes many aspects such as environmental safety, pension, residents' health, basic medical care, and convenient service. Therefore, these factors should be fully considered in the setting process of the evaluation system, so as to make the evaluation system holistic.
- (4) Principle of regional differences [18]: the design of the evaluation system should fully consider the different needs of different regions, For example, some communities where older people live should consider community healthcare and community medical care, while white-collar communities may

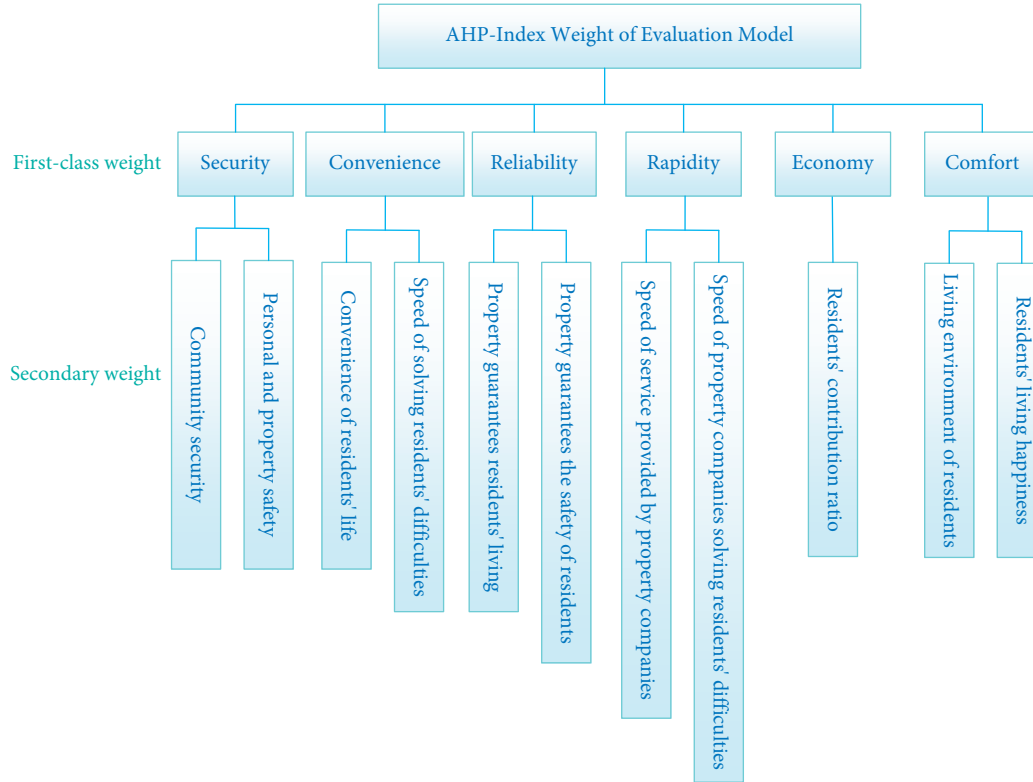


FIGURE 1: Weight relation diagram of primary and secondary levels.

pay more attention to community health and community entertainment and should consider the impact of outsiders on a series of public infrastructure enjoyed by the community after the implementation of the block system.

2.2.2. AHP Index Weight of the Evaluation Model. This model is divided into first-class weights and second-class weights, with 6 first-class weights and 11 second-class weights. The first-level weight security includes two second-level weights, community security and personal and property safety, convenience includes the convenience of residents' lives and the speed of solving residents' difficulties, reliability includes property to residents' living guarantee and property to residents' safety guarantee, rapidity includes property companies' service speed and property companies' difficulty solving speed, economy mainly refers to residents' payment ratio, and comfort includes owners' living environment and owners' living happiness. Because the AHP evaluation model reflects the satisfaction index of community residents for the quality of property services as a whole, rather than an individual satisfaction index, it is necessary to avoid individual economic differences of residents. Using the proportion of payment as the second-class weight is an effective index to reflect the overall satisfaction of residents, so it is feasible. Each weight relationship is shown in Figure 1.

2.2.3. Experimental Simulation of the AHP Evaluation Model. (1) Establish a hierarchical structure model: the index

weight is set as target layer A, the first-level weight is criterion layer B, and the second-level weight is index layer C.

(2) Constructing judgment matrix: we compare several indicators at the same level that affect the satisfaction of owners with the construction of smart communities and compare the influence degree of the indicators at the upper level in pairs, thus forming a judgment matrix. After that, each index of the same layer is compared with the corresponding index of the previous layer in pairs, and finally, the relative importance of each factor of each layer is judged, respectively.

(3) Single arrangement of layers and total arrangement of layers: in this study, well-known property companies in Qingdao, such as New Age Property and Yinshengtai Property, are selected, and senior representatives of some community owners' committees are recruited to form an evaluation team to evaluate the index evaluation system of smart communities. Through in-depth interviews, the importance of each factor at the criterion level and index level is scored by the scaling method [19]. The specific scores obtained are given in Table 2.

Taking "security" as an example, the judgment matrix is constructed as B_1 :

$$B_1 = \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{9} \\ 5 & 1 & \frac{1}{3} \\ 9 & 3 & 1 \end{bmatrix}. \quad (5)$$

In the B_1 matrix, the index relationship represented by each data is as follows: $A_{12} = 1/5$: “Community security” is more important than “resident’s property security.” $A_{13} = 1/9$: the importance of “personal safety of residents” is far greater than that of “property safety of residents.” $A_{23} = 1/3$: “Community security work” is slightly more important than “property safety of owners.”

(4) Consistency test of matrices: the steps of calculation by the summation method are as follows:

Step 1: Normalize the elements of B by column [20].

Step 2: $\bar{B} = (\bar{B}_{ij}), \bar{B}_{ij} = b_{ij} / \sum_{i=1}^n b_{ij}, i, j = 1, 2, L, L, n$.

Step 3: Add \bar{B} by rows to get $\bar{F} = [\bar{f}_1, \bar{f}_2 L L \bar{f}_n], \bar{f}_i = \sum_{i=1}^n \bar{b}_{ij}$.

Step 4: After normalization \bar{F} , get $F = [\bar{f}_1, \bar{f}_2 L L \bar{f}_n]T, K1 = \bar{K}i / \sum_{i=1}^n \bar{K}i$.

Table 3 provides the judgment matrix, results, and f_i results of safety in the evaluation of residents’ satisfaction with development community construction.

There is $a_{ij} = a_{ik}/a_{kj} (k = 1, 2, \dots, n)$ for any $i, j = 1, 2, \dots, n$. When the matrices are completely consistent, there is $\lambda_{\max} = n$; when there is consistency error in the judgment matrix, $\lambda_{\max} > n$. At the same time, the greater the error, the greater the value of λ_{\max} .

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n \frac{\sum_{j=1}^n a_{ij} k_j}{k_i} \quad (6)$$

The λ_{\max} value of the judgment matrix of residents on the safety index of the development community is 3.1384.

C.I is used as an index to test the consistency of the above judgment matrix:

$$C.I = \frac{\lambda_{\max} - n}{n - 1} = 0.0135. \quad (7)$$

After calculation, when $n = 3$, the correction coefficient R.I = 0.49.

Therefore, there is $C.R = C.I/RI = 0.0135/0.49 = 0.027551 < 0.1$, so it is concluded that the judgment matrix is consistent. In the same way, the values of other weights can be calculated.

When the judgment matrix of “convenience B_2 ” is $B_2 = \begin{bmatrix} 1 & 5 & 9 \\ 5 & 1 & 4 \\ 1/9 & 1/4 & 1 \end{bmatrix}$, the weight vector $F_2 = \{0.7241, 0.1883\}$.

When the judgment matrix of “reliability B_3 ” is $B_3 = \begin{bmatrix} 1 & 1/5 & 1/9 & 1/3 \\ 5 & 1 & 1/3 & 3 \\ 9 & 3 & 1 & 6 \\ 3 & 1/3 & 1/6 & 1 \end{bmatrix}$, the weight vector

$F_3 = \{0.2374, 0.4879\}$. When the judgment matrix of “rapidity B_4 ” is $B_4 = \begin{bmatrix} 1 & 3 \\ 1/3 & 1 \end{bmatrix}$, the weight vector $F_4 = \{0.2374, 0.4879\}$. When the judgment matrix of

“Economy B_5 ” is $B_5 = \begin{bmatrix} 1 & 1/3 & 5 \\ 3 & 1 & 7 \\ 1/5 & 1/7 & 1 \end{bmatrix}$, the weight vector

$F_5 = \{0.113, 0.2713\}$. When the judgment matrix of “Com-

fort B_6 ” is $B_6 = \begin{bmatrix} 1 & 4 & 8 \\ 1/4 & 1 & 7 \\ 1/8 & 1/8 & 1 \end{bmatrix}$, the weight vector

$F_6 = \{0.5328, 0.0654, 0.635, 0.3112\}$.

By the same token, the judgment matrix of the first-level index B in the AHP evaluation model is

$$B = \begin{bmatrix} 1 & 0.1358 & \frac{1}{2} & \frac{1}{5} & 0.11 & \frac{1}{3} \\ 6 & 1 & 6 & 2 & \frac{1}{3} & 3 \\ 2 & 0.2321 & 1 & 0.2 & 0.2321 & 0.2 \\ 1 & 4 & 5 & 1 & \frac{1}{5} & 3 \\ \frac{1}{4} & 1 & 7 & 5 & 1 & 5 \\ 3 & \frac{1}{4} & 5 & \frac{1}{3} & \frac{1}{5} & 1 \end{bmatrix}, \quad (8)$$

$$F = \{0.0283, 0.2496, 0.0325, 0.4289, 0.0831\}.$$

From this, the values of the weights of each level of the final evaluation model can be obtained, as given in Table 4.

2.2.4. Experimental Conclusions of the AHP Evaluation Model.

From the AHP evaluation model, we can see that in the future open community planning, the security and convenience of the community are extremely important. First of all, the safety management of the community should be strengthened, but more cameras must be placed to check suspicious things, so as to ensure that there is no dead angle in the photography field of vision and achieve the effect that the security system can effectively meet the development needs. Second, a set of risk prevention measures should be scientifically formulated to avoid the occurrence of sudden dangerous events.

Due to the serious aging trend in today’s society, the proportion of children and children is increasing, and the services for the life, health, and personal safety of the elderly and children are also extremely important, so the community residents’ service health platform should be added. It can provide physical examination services or medical services for a certain period of time and can also install some health monitoring instruments, such as the pulse test, blood sugar test, and body fat measurement, and can provide residents with physical fitness tests free of charge at any time.

In the future block planning work, in the community construction module, we should first meet the needs of residents and take the residents’ true happiness experience as

TABLE 2: Scores of each evaluation index.

Criterion level	Score	Index level	Score
Safety	9.1	Community security	8.7
		Personal safety of residents	8.9
		Residents' property safety	8.1
Convenience	8.5	Convenience of residents' life	7.5
		Degree of difficulty resolution	6.1
Reliability	7.6	Property safety for residents	8.3
		Property guarantees for residents' living services	7.8
Rapidity	6.3	Speed of life service provided by property	6.5
		The speed of solving the difficulties of residents by property	7.1
Economy	7.1	Proportion of residents' contributions	7.3
		Quality of property services	8.6
Comforts	8.6	Community greening environment	7.8
		Community safety environment	8.8
		Residents' happiness index	7.1
		Home care	8.7

TABLE 3: "Security" weight set.

B_1	C_1	C_2	C_3	K_i
C_1	1	0.3	0.1135	0.0748
C_2	5	1	1/3	0.2563
C_3	9	3	1	0.6754

TABLE 4: Evaluation weight of residents' satisfaction with open community.

Weight value	Safety 0.4236	Convenience 0.0831	Reliability 0.1349	Rapidity 0.0283	Economy 0.2496	Comfort 0.0325	Total weight value of C layer sorting	Importance ranking
Community security	0.0597	0	0	0	0	0	0.2894	2
Personal safety of residents	0.3214	0	0	0	0	0	0.3657	1
Residents' property safety	0.6950	0	0	0	0	0	0.1357	4
Convenience of residents' life	0	0.7241	0	0	0	0	0.2647	3
Degree of difficulty resolution	0	0.1883	0	0	0	0	0.0134	15
Property security for residents	0	0	0.0532	0	0	0	0.0754	6
Residents' living service guarantee	0	0	0.0401	0	0	0	0.0271	9
Life service speed provided by property	0	0	0	0.2374	0	0	0.0134	14
The speed of solving the difficulties of residents by property	0	0	0	0.4879	0	0	0.0548	8
Residents' contribution ratio	0	0	0	0	0.1130	0	0.0225	1
Property service quality	0	0	0	0	0.2713	0	0.0628	7
Community greening environment	0	0	0	0	0	0.5328	0.0040	12

the core essence. From the community life, environment, and other aspects of transformation, earnestly do a good job in the basic maintenance and transformation of the community.

3. Protection and Planning of Cultural Tourism Blocks

At present, the commercialization of historical sites is gradually becoming a status quo. Among them, most famous cultural blocks take tourist flow and tourist activity track as one of the main local economic sources. This situation makes

the original sense of history greatly reduced, but the commercial money atmosphere is getting hotter and hotter, destroying the original antique cultural atmosphere. Therefore, we need to use the emerging Internet of Things technology to plan cultural tourism blocks on the premise of protection.

3.1. Problems. First of all, regional culture and commercialization are mixed. In cultural areas, all kinds of goods are sold, either in the form of stalls or in the form of shops set up with blocks as the background. Moreover, its products are various and mixed, and at the same time, the products have

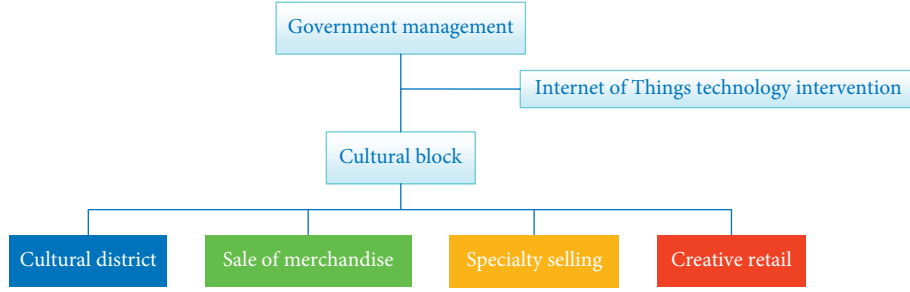


FIGURE 2: Functional partition.

the attribute of gorgeous and bright colors, which easily causes tourists' aesthetic fatigue under the interweaving of various colors, thus leading to tourists' visual deviation of ancient buildings [21]. At the same time, vendors' cries are loud and messy, which will greatly damage the cultural atmosphere of the block.

Second, the demand for goods with cultural characteristics is not high. In the commercial area, catering and local specialties have become the hottest economic industries, while the goods related to local culture are widely sold on major network platforms, which makes them lose the tourism demand that they must buy in their areas, thus leading to the serious situation that the commercialization model is biased in some aspects.

3.2. Solution. First of all, functional zoning is carried out, which divides the cultural area from the commercial area, and at the same time, it is subdivided according to the types of goods sold. Second, the Internet of Things technology is implanted, and the Internet of Things is used to promote the mutual communication between various business styles in culture, so as to achieve the purpose of common progress. It can also realize the tripartite interaction among cultural products, tourists, and businesses, thus effectively realizing the reuse of cultural carriers and realizing the revitalization of cultural functions in the process of cultural blocks. You can also use the Internet to build the connection with equipment and machines. For example, when the sensor detects that the decibel value is too high, it automatically sends information to each automatic noise reduction system. The system automatically identifies the noise area and turns on the noise reduction function at a special fixed point. The general flowchart of the idea is shown in Figure 2.

4. Optimization of the 3-Block Road Network

When residents face travel demand, they will have a variety of routes to choose. At the same time, the economic benefits and cost losses of each route are different, such as the environmental pollution value caused by vehicle exhaust, the consumption cost of vehicles, and the time value spent. Therefore, the method of using Internet of Things technology to study a route with low cost and high economic benefits is considerable.

4.1. Residents' Travel Route Choice Behavior. With the development of modern intelligent transportation technology based on Internet of Things technology, travelers can obtain the driving cost of n routes between starting and ending points through navigation tools (such as Baidu map, Tencent map, and Gaode map) and choose the route with lower cost according to preferences (shortest distance, shortest free flow time, and decision inertia). Before the opening of the closed block, because of its large size, urban travelers have a long detour distance, which is easy to produce the phenomenon of traffic superposition near it. After the block is opened, the roads inside the block will be made public, and travelers will have more choices, thus relieving the flow pressure of traffic roads in the surrounding areas. However, when urban travelers enter the block, they may face pedestrians crossing the street and motor vehicles stopping at the roadside. Therefore, when the difference between the travel cost across the block and the travel cost without crossing the block is within its acceptable range, travelers often refuse to choose the travel route across the block, that is, there is travel inertia.

4.2. Building a Bilevel Programming Model

4.2.1. Lower Model. The travel route of residents in the block will definitely cross the block. If the flow of internal travelers on a certain path is positive, the travel cost of this path is less than that of other unused paths crossing blocks, and its complementary conditions can be expressed as

$$f_{ri,mi}^w \cdot (C_{ri}^w - \mu_i^w) = 0, C_{ri}^w - \mu_i^w \geq 0, \quad \forall ri \in R_i^w, w \in W. \quad (9)$$

where mi means the residents of the block, $f_{ri,mi}^w$ is the flow of travelers inside ri on the path across blocks, C_{ri}^w is the travel cost of path ri , μ_i^w is the minimum travel cost of crossing blocks between the OD pair w , and R_i^w is the set of paths across blocks between OD pairs w .

For urban travelers, if the flow of urban travelers on a route aa crossing a block is positive, the travel cost of this route is less than that of other unused routes, and the complementary conditions are as follows:

$$f_{ri,m0}^w \cdot (C_{ri}^w - \mu_r^w) = 0, C_{ri}^w - \mu_r^w \geq 0, \forall ri \in R_i^w, \quad \forall r \in R^w, w \in W, \quad (10)$$

where m_0 means the city travelers, f_{r_i, m_0}^w is the traffic of urban travelers on the path r_i , μ_r^w is the optimal travel cost between OD and w , R^w is the set of all paths between OD pairs w .

If the flow of urban travelers on a path r_0 that does not cross the block is positive, then for the urban travelers on the path, the travel cost of the path is less than the travel cost of other unused paths r_i that do not cross the block, and the difference between the travel cost and the path within the block is less than the acceptable threshold ε_w of the travelers who choose the path, which can be expressed as

$$\begin{aligned} C_{r_0}^w - \mu_0^w \geq 0, C_{r_0}^w \mu_i^w \geq \varepsilon_w, \\ \text{if } f_{r_0, m_0}^w = 0, \quad \forall r_0 \in R_0^w, \forall r_i \in R_i^w, w \in W. \end{aligned} \quad (11)$$

$$\begin{aligned} C_{r_0}^w - \mu_0^w \geq 0, C_{r_0}^w \mu_i^w \geq \varepsilon_w, \\ \text{if } f_{r_0, m_0}^w = 0, \quad \forall r_0 \in R_0^w, \forall r_i \in R_i^w, w \in W, \end{aligned} \quad (12)$$

where f_{r_0, m_0}^w is the flow of urban travelers on the path r_0 , $C_{r_0}^w$ is the travel cost of path r_0 , μ_0^w is the minimum travel cost

between the OD pair w that does not cross the block, R_0^w is the set of paths between OD pairs w that do not cross super blocks.

Formulas (11) and (12) are expressed as complementary conditions, and the transfer function is introduced [22]:

$$\gamma_1(\psi, \Gamma, \Lambda) = \begin{cases} 0, & \text{if } \psi = 0 \text{ \& } 0 \leq \Gamma \leq \Lambda, \\ \psi + \Gamma, & \text{other,} \end{cases} \quad (13)$$

where ψ, Γ, Λ is an independent waste negative variable, which can be expressed in the following complementary form based on equations (11), (12), and (14):

$$f_{r_i, m_0}^w \cdot [\gamma_1(C_{r_0}^w - \mu_0^w, C_{r_0}^w - \mu_i^w, \varepsilon_w)] = 0, \quad (14)$$

where $\gamma_1(C_{r_0}^w - \mu_0^w, C_{r_0}^w - \mu_i^w, \varepsilon_w) \geq 0, \forall r_i \in R_i^w, w \in W$.

The above complementary conditions (9), (10), and (14) can be extended to a mixed equilibrium model, which is expressed by the following mathematical programming model:

$$\begin{aligned} \min F = & (f_{r_i, m_i}^w, f_{r_i, m_o}^w, f_{r_o, m_o}^w, \varepsilon_w) = \sum_{w \in W} \sum_{r_i \in R_i^w} f_{r_i, m_i}^w \cdot (C_{r_i}^w - \mu_r^w) + \sum_{w \in W} \sum_{r_i \in R_i^w} f_{r_i, m_o}^w \cdot (C_{r_i}^w - \mu_r^w) \\ & + \sum_{w \in W} \sum_{r_o \in R_o^w} f_{r_o, m_o}^w \cdot [\gamma_1(C_{r_0}^w - \mu_0^w, C_{r_0}^w - \mu_i^w, \varepsilon_w)], \\ & \sum_{w \in W} \sum_{r_i \in R_i^w} \delta_{r_i, a}^w \cdot f_{r_i, m_i}^w + \sum_{w \in W} \sum_{r_i \in R_i^w} \delta_{r_i, a}^w \cdot f_{r_i, m_o}^w + \sum_{w \in W} \sum_{r_o \in R_o^w} \delta_{r_o, a}^w \cdot f_{r_o, m_o}^w = x_a, \quad \forall a \in A, \\ & \sum_{r_i \in R_i^w} f_{r_i, m_i}^w + \sum_{r_i \in R_i^w} f_{r_i, m_o}^w + \sum_{r_o \in R_o^w} f_{r_o, m_o}^w = d_{m_i}^w + d_{m_o}^w \cdot \sum_{r_i \in R_i^w} P_{r_i, m_o} + d_{m_o}^w \cdot \sum_{r_o \in R_o^w} f_{r_o, m_o} = d_w, \\ & P_{r_i, m_o} = P\{r_i | C_{k_o}^w - C_{r_i}^w \geq \varepsilon^w\} + P\{r_i | C_{k_o}^w - C_{r_i}^w \leq \varepsilon^w\} \cdot \tau, \quad \forall k_o \in K_o, r_i \in R_i, \\ & P_{r_o, m_o} = P\{r_o | C_{r_o}^w - C_r^w \leq 0\} + P\{r_i | C_{r_o}^w - C_{k_i}^w \leq \varepsilon^w\} \cdot (1 - \tau), \quad \forall r_o \in R_o, \forall k_i \in R_i, r \in R, \\ & C_{r_i}^w - \mu_i^w \geq 0, \quad \forall k_i \in R_i^w, w \in W, \\ & C_{r_i}^w - \mu_i^w \geq 0, \forall k_i \in R_i^w, \quad \forall r \in R^w, w \in W, \\ & \gamma_1(C_{r_0}^w - \mu_0^w, C_{r_0}^w - \mu_i^w, \varepsilon_w) \geq 0, \quad \forall r_i \in R_i^w, \forall r_o \in R_o^w, w \in W. \end{aligned} \quad (15)$$

In the model, F represents the optimization function, which is used to obtain the overall minimum distance between block resident travelers and city travelers under complementary conditions (9), (10), and (14), that is, the sum of the distance between block resident travelers and equilibrium points and the distance between city travelers and equilibrium points is minimum.

4.2.2. Upper Layer Model. In order to ensure that the optimized route of the block road network is more in line with the actual needs, this study constructs a planning model with the minimum total cost of the system under the conditions of multiple factors (travel cost, environmental factors, and potential safety hazards) as follows.

(1) Optimize the objective function: ignoring other factors, it is approximately considered that the travel cost is equal to the sum of the travel time cost and the travel cost itself, and the calculation formula is as follows [23]:

$$\begin{aligned} \min Tc = & \left(\lambda \cdot \sum_{a_0 \in A_0} ta_0 \cdot xa_0 + \sum_{a_0 \in A_0} la_0 \cdot xa_0 \cdot ba_0 \right) \\ & + \left(\lambda \cdot \sum_{a_i \in A_i} tai \cdot xaii + \sum_{a_i \in A_i} tai \cdot xaii \cdot bai \right), \end{aligned} \quad (16)$$

where Tc is the total cost of road network driving, λ is the traveler's time value (VOT) parameter, ta_0 and tai are the vehicle travel times of road sections a_0 and a_i , la_0 and lai are

the length of road sections, $xa0$ and xai are the road section flows, $ba0$ and bai are the unit operating expenses of cars, $A0$ is the collection of road sections outside the block, and Ai is the collection of sections within the block.

When the block is opened, the vehicle circulation inside it is more frequent, but the exhaust pollution emitted by a large number of vehicles during driving will seriously affect the residents and their surrounding environment. Therefore, when making decisions, it is necessary to bring the environmental pollution loss in the block into the optimization goal. The specific calculation formula is as follows:

$$\min Ec = \sum_{a0 \in A0} Eai \cdot Qc, \quad (17)$$

where Ec is the cost of environmental pollution loss, Qc is the average loss cost per unit pollutant (including tail gas treatment cost and health loss cost), and Qc is the unit road pollutant emissions; this study focuses on the calculation of the main pollutants (CO, HC, and NOX) produced when the vehicle exhaust emissions. The specific calculation formula is as follows [24]:

$$Eai = \sum_{k \in K} EFk \cdot xailai, \quad (18)$$

where EFk is the emission factor of pollutant k , and $EFco = 0.5216\gamma_{co}$, $EFHC = 0.0634\gamma_{HC}$, $EFNOx = 0.0198\gamma_{NOx}$, and γ_{CO} , γ_{HC} , and γ_{NOx} are calculated as the velocity correction factors, which can be obtained from Table 5.

Existing studies have shown that safety is the focus of aborigines' concerns about block opening, so this study takes the possible losses caused by "potential" traffic accidents into consideration. The calculation formula is as follows [25]:

$$\min ARc = \sum_{ai \in Ai} 0.0035 \cdot I_{ai}^{0.6724} \cdot X_{ai}^{0.9679} \cdot Sc, \quad (19)$$

where ARc is the loss cost of potential safety hazards, Sc represents the average cost of property damage per accident (including direct property damage and indirect property damage). It should be pointed out that the average period of each traffic accident is often long, so it is necessary to evenly allocate the cost to the same statistical time interval as other optimization objectives in actual calculation.

The above problems are transformed into a single-objective optimization model with the lowest cost, which is as follows.

(2) Decision variables: in this study, the decision variable of the upper-level planning model is the decision variable yai whether the road section is open or not:

$$yai = 0, 1ai \in Ai. \quad (20)$$

In the formula, $yai = 0$ means that road section ai is not open; $yai = 1$ stands for section ai open in both directions.

(3) Constraints: the important purpose of opening blocks is to relieve the traffic pressure of urban main roads, but the traffic conditions inside blocks cannot be ignored, so the saturation of internal sections after opening blocks should be within the limit of maximum expected saturation.

TABLE 5: Comparison table of speed correction factor values of small vehicles.

Pollutant	Velocity interval (km/h)				
	<20	20-30	30-40	40-80	>80
CO	1.69	1.26	0.79	0.39	0.62
HC	1.68	1.25	0.78	0.32	0.59
NO _x	1.38	1.13	0.90	0.86	0.96

$$Sai \leq \bar{Sai} ai \in Ai, \quad (21)$$

where Sai is the saturation of the road section inside the block and \bar{Sai} is the maximum expected saturation.

4.3. Experimental Comparison. Referring to the average block size of cities in China, the longest horizontal length of the block is set to 2.1 km and the longest longitudinal length is set to 1.8 km. Its external road network includes 34 intersection nodes, 38 two-way trunk roads, 26 two-way secondary trunk roads, and 11 urban branch roads. The interior of the block includes 9 intersections and 42 two-way branch roads. Taking Chengdu as an example, based on the actual data of the city, the parameters such as traffic capacity, driving speed, and travel demand of the example road section are set as follows: the single-side traffic capacity of the urban main roads in this area is 2200 Veh/h, and the speed is 40 km/h; the single-side traffic capacity of the secondary trunk road is 1800 Veh/h and the speed is 30 km/h, while the single-side traffic capacity of the branch road is 800 Veh/h and the speed is 30 km/h.

Suppose: when the traveler has low inertia decision, $\varphi = 0.2, \tau = 0.8$; when the traveler has a low inertia decision, $\varphi = 0.35, \tau = 0.65$; when the traveler has moderate inertia decision, $\varphi = 0.5, \tau = 0.5$; when the traveler has high inertia decision, $\varphi = 0.65, \tau = 0.35$; when the traveler has high inertia decision, $\varphi = 0.8, \tau = 0.2$; and when the traveler is completely rational, $\varphi = 0, \tau = 1$.

With the decrease of residents' travel decision inertia, the effect of opening blocks on relieving traffic pressure in urban sections is obviously improved. For example, although the travel cost inside the block becomes higher, the travel cost outside the block becomes lower, and the total system cost also becomes lower; the total travel time of the road outside the block decreased from 106,848 min to 101,116 min, a decrease of 5.36%. The results are shown in Figures 3 and 4.

In addition, after the opening of the block, the travel time of residents in the block is on the rise as a whole, and external vehicles can enter the inner lane of the block, which affects daily travel to some extent. As a result, the traffic road exhaust emissions and potential safety hazards in the block also show a positive growth trend. Among them, the exhaust emissions inside the block increased by about 3.7 times, and the value of potential safety hazards increased by about 3.6 times. The results are shown in Figure 5.

With the opening of the block roads, the average travel time of urban residents showed a downward trend, with an overall average decline time of about 33 min and a decrease

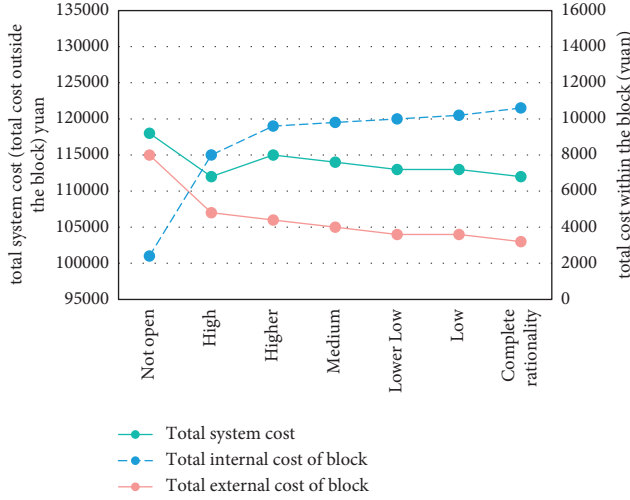


FIGURE 3: Travel cost change.

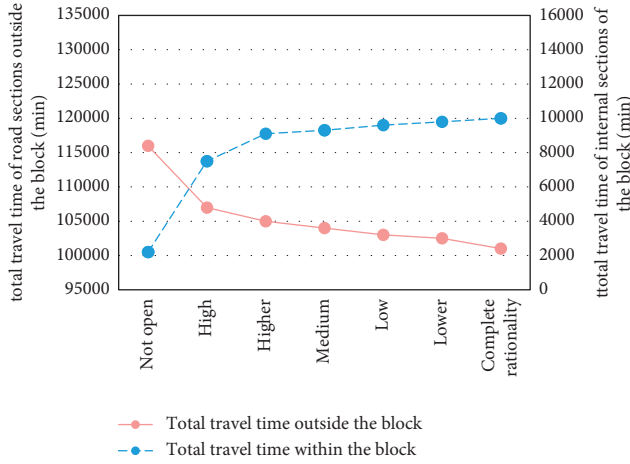


FIGURE 4: Total travel time of road network sections.

of about 12.09%. Travel conditions have improved. The results are shown in Figure 6.

The result of these data is that when the traveler's decision inertia is high, the traveler is unwilling to transfer from the current choice to the better path in the block; however, when the traveler's decision inertia is low, the inner branch road of the block can be effectively utilized. Therefore, when implementing the block opening policy, the travel decision inertia of travelers can be reduced through reasonable information induction.

4.4. Optimization Results. By referring to other relevant professional contributions and simulation verification in this study, it is assumed that the tolerance coefficient $\varphi = 0.5$, travel preference selection coefficient $\tau = 0.5$, maximum acceptable indifference threshold $\varepsilon_{\max}^w = 0.25$, traveler's travel time value $\lambda = 0.25$, automobile unit operating cost $b_a = 1$ yuan/km ($v = 40$ km/h) - $b \pm \sqrt{b^2 - 4ac}/2a$, $b_a = 1.2$ (yuan/km)/($v = 30$ km/h) = 30 km/h, environmental pollution loss cost $Q_c = 0.1$ yuan/g, and single accident loss cost

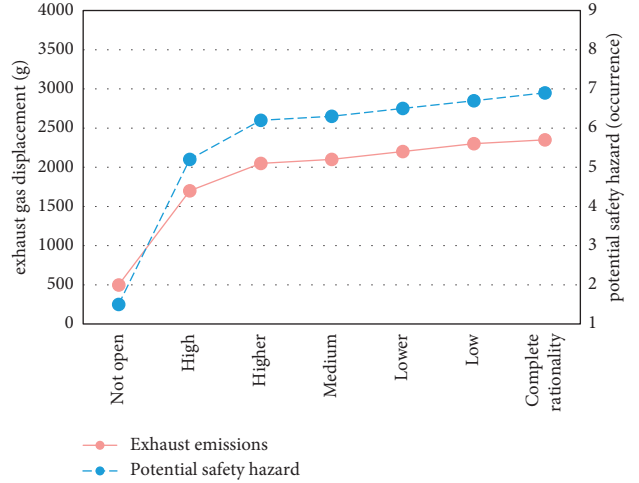


FIGURE 5: Exhaust emissions and potential safety hazards.

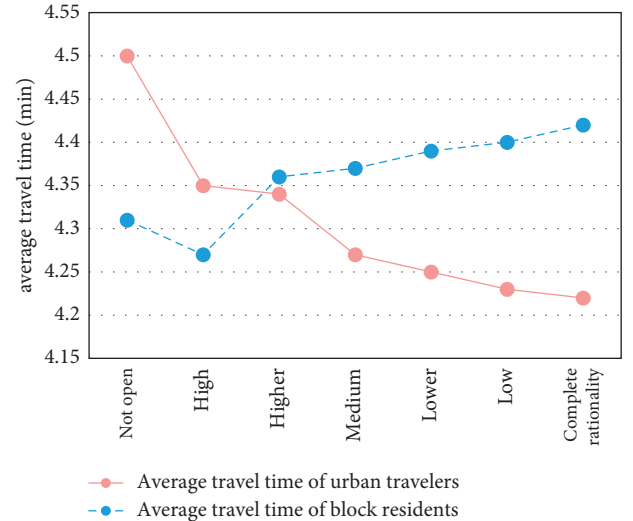


FIGURE 6: Average travel time of travelers.

$S_c = 20\bar{\pi}/h$, $\bar{S}_{a_i} = 0.9$ of urban travelers in the system. Compared with before and after the optimization of the block road network, the total cost of the system changed to 113,820 yuan, a decrease of 2.63%. The travel time of the external section of the block road network decreased significantly, from 115,871 min to 103,674 min, a decrease of 10.53%. The average saturation of urban main roads decreased from 0.92 before opening to 0.87, a decrease of 6.23%, and the average saturation of secondary roads decreased from 0.84 to 0.74, a decrease of 11.56%. The congestion of the road network outside the block was obviously improved. However, some changes have taken place in the inner travel of the block. First, the travel time has increased from the initial 2125 min to 9264 min, and the average saturation of the road section has increased from the initial 0.14 to 0.58, which is still in a basically unblocked state. In addition, the emission of automobile exhaust pollution in the block increased by 1611 g/h, and the number of potential traffic accidents increased by 4.67/year, all of which may affect the daily life of residents.

5. Conclusion

In this study, the historical blocks are roughly divided into three modules for research, which are residential communities, cultural tourism areas, and traffic routes. The complete information game model is used to analyze the imbalance between property and residents. Adopt the AHP fuzzy evaluation method to detect the current life happiness coefficient of urban residents and collect residents' opinions and suggestions for the future planning of historical blocks. The bilevel programming model is used to solve the block road network optimization scheme, and the effectiveness of the optimization results is tested. The following conclusions are drawn:

- (1) Residents have high expectations for the safety and convenience of the future community, so we should focus on the research of future block planning from these two angles
- (2) Because the aging trend of today's social situation is getting worse and even more serious in the future, it is necessary to increase the community residents' service health platform. It is required to provide physical examination services or medical services for a certain period of time. At the same time, some health monitoring instruments can be installed, such as the pulse test, blood sugar test, and body fat measurement, which can provide residents with physical fitness tests free of charge at any time.
- (3) The commercialization of cultural tourism blocks is serious, the cultural atmosphere is not strong, and the main functions are not prominent. Therefore, it is of considerable practical significance to partition functional areas. The functions are divided into cultural area, retail sales of goods, specialty sales, and creative retail. At the same time, these four areas are connected by using Internet of Things technology to achieve the purpose of information sharing, real-time monitoring, and timely handling of emergencies.
- (4) Opening the entrance and exit of the block will help relieve the pressure on the main road, but it will also have a certain impact on the daily life of residents. The overall travel time of the block will decrease, but the travel time within the block, the incidence of safety accidents, and the emission of automobile pollution will increase. The decision inertia of urban travelers, the internal demand of blocks, and the demand of urban travel will also affect the opening effect of blocks. When the travel demand within the block is low (the average saturation of the internal road network is less than 0.6 under the condition of not opening the block) and the urban transit travel demand is high (the average saturation of the external road network is greater than 0.72 and less than 1.07), the open block can effectively alleviate the congestion of the external road network without obvious impact on the internal road network of the block. The higher the decision inertia of travelers, the

higher the internal and external travel demand that the open block policy can cope with, that is, the stronger the adaptability of the policy.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] A. Kumar and D. S. Meshram, *Future of Cities: Planning, Infrastructure and Development*, Taylor and Francis, Oxfordshire, UK, 2022.
- [2] A. Wong, Y. Qiu, and G. Wang, "On the continuation of context and vitality revival of qishan uocheng historic district in yantai," *Architecture and Culture*, vol. 11, pp. 218-219, 2018.
- [3] S. Malekpour, R. R. Brown, and F. J. de Haan, "Strategic planning of urban infrastructure for environmental sustainability: Understanding the past to intervene for the future," *Cities*, vol. 46, pp. 67-75, 2015.
- [4] Y. Shen, Y. Chai, and X. Ma, "Concept, Model and architecture of People-oriented Smart Community," *Modern Urban Research*, vol. 10, pp. 13-17 +24, 2014.
- [5] J. Fu, L. Y. Hui, and X. D. Wang, "Complete Information dynamic game analysis in tacit knowledge Sharing Process within Modern Companies," *Advanced Materials Research*, vol. 421, p. 421, 2011.
- [6] R. Vladimir, P. Dinko, and K. Goran, "Operational Risk Management Using multi-criteria pssessment (AHP Model)," *Tehnički vjesnik*, vol. 28, no. 2, 2021.
- [7] J. Zhang, "Research on the strategy of building intelligent public transport system with Internet of Things technology in Huilongwo historical and cultural block of Xuzhou," *Jushe*, vol. 30, pp. 13-15 +25, 2021.
- [8] X. Liu, B. Shuai, L. Zhang, and S. Zhang, "Optimization method of super block road network considering traveler's decision inertia [J/OL]," *Journal of Transportation Engineering and Information*, pp. 1-11, 2022.
- [9] H. A. Simon, "A Behavioral Model of Rational Choice," *The Quarterly Journal of Economics*, vol. 69, no. 1, pp. 99-118, 1955.
- [10] H. Xu, H. Yang, J. Zhou, and Y. Yin, "A Route Choice Model with context-dependent Value of Time," *Transportation Science*, vol. 51, no. 2, pp. 536-548, 2017.
- [11] K. Liu and J. Zhou, "Research on mixed user equilibrium model induced by traffic information," *System Engineering Theory and Practice*, vol. 40, no. 2, pp. 415-425, 2020.
- [12] H. Xu, Y. Liu, and W. Xu, "Stochastic network path selection and traffic flow assignment model considering decision inertia," *System Engineering Theory and Practice*, vol. 41, no. 4, pp. 1010-1017, 2021.
- [13] J. Zhao and P. Sarana, "Commercialization suitability evaluation of historical and Cultural Blocks in, Beijing, China," *Journal of Service Science and Management*, vol. 14, no. 3, 2021.
- [14] W. Deng, "Research on the problems in the construction of Smart community in china and Its countermeasures," in *Proceedings of the 3rd International Conference on Economics, Management, Law and Education (EMLE 2017)(Advances in*

- Economics, Business and Management Research*, pp. 234–238, Zhengzhou, China, 2017.
- [15] B Sawalhal and T Sayed, “Evaluating safety of urban arterial roadways,” *Journal of Transportation Engineering*, vol. 127, no. 2, pp. 151–158, 2001.
- [16] Z. P. Song, “Remote video web monitoring system design of port coastal ships Based on UDP,” *Advanced Materials Research*, vol. 898, pp. 747–750, 2014.
- [17] Anonymous. Teaming, “Up with cities on building smart communities,” *Electric Perspectives*, vol. 44, no. 4, 2019.
- [18] H. Zhao, C. Yang, W. Guo, L. Zhang, D. Zhang, and H Correction Zhao, “Automatic estimation of crop disease severity levels based on vegetation index normalization,” *Remote Sensing*, vol. 12, p. 1930, 2020.
- [19] Y. Jia, H. Han, and J. Ma, “Pension service system based on grey prediction model,” *International Journal of Computational and Engineering*, vol. 5, no. 1, 2020.
- [20] M L Kappers Astrid and B Schakel Wouter, “Comparison of the haptic and visual deviations in a parallelity task,” *Experimental brain research*, vol. 208, no. 3, 2011.
- [21] J. Zhang, “Exploration and research on functional planning of commercial areas in historical and cultural blocks based on Internet of Things technology-taking Huilongwo historical and cultural block in Xuzhou as an example,” *Cultural Industry*, vol. 3, pp. 147–150, 2021.
- [22] W. Y. Szeto and H. K. LO, “Dynamic traffic assignment: properties and extensions,” *Transportmetrica*, vol. 2, no. 1, pp. 31–52, 2006.
- [23] A. Irman, Y. Muharni, and Y. Andri, “Design of warehouse model with dedicated policy to minimize total travel costs: a case study in a construction workshop,” *IOP Conference Series: Materials Science and Engineering*, vol. 909, no. 1, 2020.
- [24] M. Duell, H. Grzybowska, D. Rey, and S. T. Waller, “Strategic dynamic traffic assignment incorporating travel demand uncertainty,” *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, 2019.
- [25] M. Martcheva, N. Tuncer, and C. N. Ngonghala, “Effects of social-distancing on infectious disease dynamics: an evolutionary game theory and economic perspective,” *Journal of Biological Dynamics*, vol. 15, no. 1, pp. 342–366, 2021.

Research Article

MOOC Teaching Model of Basic Education Based on Fuzzy Decision Tree Algorithm

Zhang Yuanyuan 

College of Education, Taiyuan Normal University, Taiyuan, Shanxi 030619, China

Correspondence should be addressed to Zhang Yuanyuan; zhangyy@tynu.edu.cn

Received 4 March 2022; Revised 13 April 2022; Accepted 16 April 2022; Published 8 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Zhang Yuanyuan. 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 recent years, the development of science and technology in China has greatly affected people's ways of entertainment. In the traditional industrial model, new industries and Internet industries represented by the Internet have emerged, and the Internet video business is an emerging business that has been gradually emerging in the Internet industry in recent years. Moreover, this new teaching method has been gradually noticed in simple education, such as MOOC, I want to self-study network, and Smart Tree, and other online learning websites have sprung up. At present, the epidemic environment makes people pay more attention to this convenient and wide range of online video education. Therefore, we need to evaluate this kind of online video teaching model from the effectiveness of this kind of method and the quality of user experience. This paper takes this as the starting point and chooses the earliest online video platform, MOOC, as the model to establish a set of perfect user experience quality evaluation methods suitable for domestic online video education mode. Considering the data source, the accuracy of the results, and other factors, we chose the industry-leading platform MOOC network as an example. Through the exploration of the MOOC teaching mode in basic education, a member experience evaluation model is established based on fuzzy decision tree algorithm. The experimental results show that the model has high accuracy and high reliability.

1. Introduction

With the popularity of online video software of Tencent, Youku, and Iqiyi, people gradually pay attention to the field of online video. The sudden epidemic has also made more and more people pay attention to the teaching model of online video education (Li et al.) [1]. Under the current epidemic situation, more and more schools choose to upload courses to online teaching platforms, which are funded by students to watch, study, and download. In addition, for some students who want to acquire professional knowledge in addition to school, some course videos taught by university teachers are the best way to learn, and video learning has become one of the ways and methods of national learning (Guand He) [2]. Therefore, in recent years, the Internet video business has gradually become an emerging business. More and more Internet companies begin to focus on and tend to the development of their video business. Development is bound to be accompanied by competition,

and the competition of online video education is becoming more and more intense. The richness of video content and video quality determine the retention and experience of members (Farnaz et al.) [3]. However, live video education evolved from the traditional video industry, but it is quite different from the traditional new media business in terms of video content, video architecture, user needs, member audience, and so on (Hu et al.) [4]. Therefore, we cannot apply the traditional video model evaluation methods to the modern video education model, so we lack some evaluation and deep understanding of the current video education model. Although there are mature platforms abroad, such as courses, we need to consider the differences among China and foreign countries in international background and industrial development environment and user behavior. We can only learn from foreign research results, but we cannot take them as a guide (Vijay et al.) [5]. Therefore, how to establish a set of user experience quality evaluation methods for China's modern online video education model is very

important and necessary. In view of this problem, we found through research that the MOOC platform is one of the earliest platforms for online video courses and video education in China. Compared with other platforms, there are hundreds of 985/211/double first-class universities on the MOOC platform, and their authority and richness of courses are far higher than other online video websites. Coupled with the influence of its years, the Muke platform can become the earliest representative model of online video education in China (fuzzy et al.) [6]. Therefore, taking Muke as an example, this paper discusses the user experience quality evaluation method of the Internet video education model.

The data cited below are from the real access data of the platform. Based on the collected data, we establish the user experience quality evaluation model of online Internet video learning through a fuzzy decision tree algorithm. For the collected data, we first clean the data and sort the collected initial text data into quantitative visual data, which is convenient for the next data analysis. After preprocessing the original data, the original data can show the login IP address and user agent field information to infer the type of access device. Finally, we get a series of data information such as member access video program information, member login information, and so on. On this basis, we use the mathematical algorithm, namely the fuzzy decision tree algorithm to statistically analyze the set to mine the characteristic laws of the data. For example, we can explore the distribution of video quality and the correlation between video quality and member retention rate, and lay a certain foundation for the next modeling work according to the mathematical characteristic laws we mine. Thus, an online video user experience quality evaluation model suitable for the current situation of Internet development in China is established, and the model is modeled based on a fuzzy number decision algorithm to evaluate and verify the accuracy and effectiveness of the model. Finally, the online video learning mode represented by the MOOC is evaluated by a fuzzy decision tree algorithm.

This paper establishes a set of perfect user experience quality evaluation methods suitable for domestic online video education mode. The research and innovation contributions include: through the exploration of MOOC teaching mode of basic education, a member experience evaluation model based on a fuzzy decision tree algorithm is established. The experimental results show that the model has high accuracy and high reliability. This set of evaluation teaching models is suitable for China's industrial development and combined with China's national conditions and provides a reference value.

2. Related Work

The evaluation of the online video learning mode is mainly to model and test the user experience under this mode. Therefore, this paper mainly explores the video viewing law and user experience of members using MOOC online video learning mode based on a fuzzy number decision algorithm. Among them, fuzzy data measurement algorithm is mainly used in video traffic measurement. Data analysis and sorting

play an important role in the cleaning process (Sun) [7]. Firstly, the video measurement method is divided into the active and passive measurements. As the name suggests, the measured party actively obtains the specific browsing data of the measured party by sending data packets, so that the browsing records and interested video content of the measured party can be obtained quickly, conveniently, and directly. Therefore, the active measurement method has strong operability and simple, flexible, and direct operation (Teekaraman et al.) [8]. Passive measurement does not need to actively send data packets back to the user but directly pulls the required data from the network server through the browsing records of the network and the network characteristics on the statistical link of the background data packets. The alternative method does not need to send data. Therefore, it does not need to occupy the network transmission space and has little impact on the stability of the network. And because the data are directly pulled out in the background, the accuracy is also higher than the former (Fuzzy D et al. 2020) [9]. In the process of specific experimental design, we often choose different measurement methods according to the needs of specific conditions. After determining the way of data acquisition, we need to sort out and count the data to explore the law of member behavior behind the data. This law is mainly the law of the whole process after the member starts to execute the viewing behavior channel, and the process ends after the closing operation, including the relationship between the frequency of member switching video, viewing time, and the distribution of member resolution characteristics. Foreign research teams have conducted research on the frequency behavior of members' video conversions. This research analyzes the frequency distribution of member visits; that is, the frequency of collected member visits and online video teaching platform is drawn as zero, one, and more times (Raharja) [10]. According to the proportion of times of each frequency, it is found that the retention rate of members who use the PC terminal to initiate online video access requests is higher than that of members on the mobile terminal or iPad terminal, and the video conversion rate is lower in the viewing process, that is, less than 5% of users have switched videos in the viewing process (Jan et al.) [11]. The author of this article further clustered the video conversion time initiated by members and found that most of the video conversion time was when the frequency began to play (Mu et al.) [12]. Finally, according to the above collected data and the law of analysis from the data, we can do the user experience research and evaluation model of the online video education mode. The MOOC teaching evaluation model for basic education we proposed is an evaluation method that is in line with our current industrial development status and can effectively evaluate the online video teaching model, taking MOOC as an example. This method draws lessons from the existing evaluation models abroad to a certain extent. Florin Dobrain and others are researchers who have accessed data analyses through real online video platforms for the first time. They mainly explore the correlation between the member service scene and user participation (Shi and Huang) [13]. This is the first experiment using real scene

data analyses, so the experimental results have certain reliability. However, the disadvantage is that the analysis conclusion of this study only depends on the actual data, and the specific reliability and accuracy have not been verified by relevant experiments. Subsequently, S. Shunmuga Krishnan et al. proposed to conduct research through quasi-experiment for the first time. They explored the relationship between video quality and member access behavior through experiments and finally found that there was a causal relationship between food quality and member access behavior (Khazali et al.) [14]. The authenticity and reliability of experimental results are more accurate and have higher credibility. As for the specific experimental methods, we can see the research on the relationship between food quality indicators and member participation in the visit process. Although s. shunmuga Krishnan et al. creatively improved the reliability of the experiment by using the experimental method in the research of the online video teaching model. At this time, the experiment mainly adopts qualitative measurement, which has a new overall impact on the correlation between the two in the experimental process, exploring the heterogeneity of scale and the rationality of the method. Subjective avoidance in the process of experimental design are not perfect and needs to be further discussed (Zhenget al.) [15].

3. Materials and Methods

The main purpose of this paper is to find a user experience evaluation scheme system suitable for the MOOC teaching mode of basic education. Firstly, the system needs to better fit the existing data on the MOOC teaching platform of basic education so that the evaluation system can become an effective proof of the credibility of the evaluation of member interview experience in the future. Therefore, the essence of the system is to reflect the correlation between video quality and member experience quality in the MOOC teaching mode of basic education and the optimal solution of linear relationship fitting. There is a certain balance between ordinary users and member users, which can neither make the ordinary video too simple nor let the member video monopolize. In the process of finding the optimal solution, it is inevitable to use data mining algorithm. Therefore, we first investigated several commonly used classification algorithms in data mining, such as the naive Bayes algorithm, decision tree algorithm, support vector machine, and the fuzzy decision tree algorithm. However, the Naive Bayes algorithm is suitable for scenes where data sets are independent of each other. However, the premise of this paper is that there is a correlation hypothesis between video quality and member access quality experience, so the classification effect of the model established by the naive Bayesian algorithm is not ideal. The decision tree algorithm is suitable for data sources that are discrete value sample sets, but the data sources and subsequent classifications we deal with regard to samples as continuous value sets, so the decision tree algorithm is not suitable for this paper. However, because the decision tree is a gradual algorithm based on dichotomy, the comprehensiveness and theory of the algorithm are

relatively simple. For the deep mining algorithm suitable for the optimal solution in this paper, the decision tree algorithm can better fit the data results by removing the differences of data sources. Therefore, this paper adopts the fuzzy decision tree algorithm to solve this problem, which is more in line with the cognitive formula of data attribute characteristics and uncertainty, in theory. In the following practical experiments, we also take the decision tree algorithm as the comparison algorithm to evaluate the accuracy of the model. Because the remaining support vector machine algorithms are suitable for small-scale sample data mining problems, there are problems of long modeling times and complex algorithms in this paper, resulting in weak practical operability and eventually elimination. Therefore, this paper finally improves on the basis of the decision tree and adopts the fuzzy decision tree algorithm with more reliable and accurate results Figure 1.

Next, we collect and preprocess the data. The data we use are all from the actual access data of the MOOC network, and their content is only the log files in the process of member access, which do not involve the specific privacy of members. MOOC is an online video learning platform for basic education jointly developed by love course and Netease cloud classroom. It can be regarded as one of the first batches of the online video teaching platforms developed in China. Its platform has been settled in hundreds of universities, including 985/211/double first-class universities such as Tsinghua University, Peking University, Nankai University, and Fudan University. During the epidemic period, more and more schools choose to upload teaching videos to the MOOC platform, where students attend classes at home and complete homework and exams.

In addition, more and more people choose the MOOC platform for direct learning when they choose extracurricular knowledge other than their major, and nonstudents want to study courses on the university campus. Therefore, the MOOC platform has the amount of data that meet the needs of experimental research. In addition, the MOOC has the technical support of Internet head and huge database and video supports network TV, IP mobile phone iPad, PC, and other playback platforms, which also provide basic data support for us to study the behavior of members accessed by different terminals. We collect the key information from the collected original data, clean the duplicate information and fill in the vacancy information, identify the access behavior, determine the key video quality, and calculate the address and location of member access. Finally, the data result we sorted out is the collection of behavior records of the MOOC online video education platform Figure 2. Each record includes three parts: video information accessed by members, identity information of members, and specific behavior process information.

Next, we will rely on the fuzzy decision tree algorithm to establish an evaluation model that reflects the correlation between the video quality visited by members and the actual experience quality of members. The following is the theoretical basis of the fuzzy decision tree algorithm based on the classification fuzziness. First, we

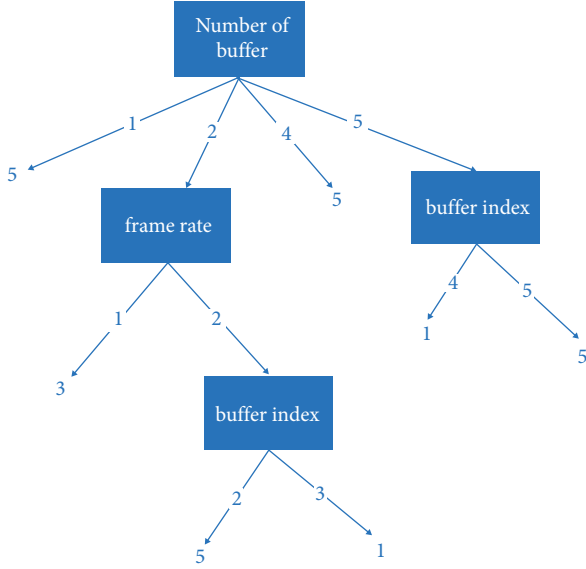


FIGURE 1: Example of a fuzzy decision tree generated by the algorithm.

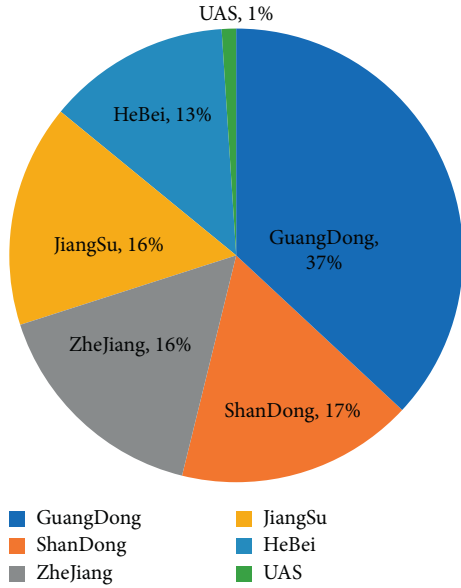


FIGURE 2: Member login area distribution statistics.

define the size of fuzzy set A , as shown in the following formula:

$$M(A) = \sum_{u \in U} \mu_A(u). \quad (1)$$

When A represents a fuzzy set on the full data set U , it is included in the function μ_A . When U is a discrete set, the fuzziness of fuzzy set U is defined in the following formula:

$$M_V(A) = -\frac{1}{m} \sum_{i=1}^m (\mu_i \ln \mu_i + (1 - \mu_i) \ln (1 - \mu_i)). \quad (2)$$

For the discrete set X , according to the fuzzy decision tree algorithm, if there is a normalized distribution of

variable Y on its set, the fuzziness of variable Y is defined as the following formula:

$$E_a(Y) = g(\pi) = \sum_{i=1}^n (\pi_i^* - \pi_{i+1}^*) \ln i. \quad (3)$$

It can be deduced from the above definition that when the variable Y can only take one value, the value of the fuzziness set is 0, which means that the variable Y does not have fuzziness at this time. When the possibility that variable Y can take any value in set X is 1, Y has the greatest fuzziness. When the numerical attribute is continuous, variable A can form a set containing S discrete semantics after fuzzy processing, in which each semantic item is a fuzzy set. In order to measure the fuzziness of A the membership function value of any variable in the fuzzy set is set u , and the fuzziness possibility distribution of continuous attribute A can be obtained. First, normalize the probability distribution of set u , as shown in the following formula:

$$\pi_{T_s}(u_i) = \frac{\mu_{T_s}(u_i)}{\max_{1 \leq j \leq S} \{\mu_{T_s}(u_j)\}}, \quad s = 1, 2, 3, \dots, S,$$

$$E_\alpha(A(u_i)) = g(\pi_T(u_i)), \quad (4)$$

$$E_\alpha(A) = \frac{1}{m} \sum_{i=1}^m E_\alpha(A(u_i)).$$

From the above, the fuzziness measurement formula of continuous variable A can be obtained, and the result can also be applied to the measurement of fuzziness of classification results. Next, we discuss the fuzziness rule and its confidence level. In the fuzziness rule, two conditional fuzzy sets A and B have been defined, and it is assumed that there is a corresponding relationship between them. Therefore, we need to define the authenticity of the rule by using the concept of confidence level, expressed in $S(A, B)$. See formula (5) for the specific calculation formula. At this time, the category possibility judgment formula of attribute variables is the following formula:

$$S(A, B) = \frac{M(A \cap B)}{M(A)} \quad (5)$$

$$= \frac{\sum_{u \in U} \min(\mu_A(u), \mu_B(u))}{\sum_{u \in U} \mu_A(u)},$$

$$\pi(C_t|E) = \frac{S(E, C_t)}{\max S(E, C_j)}, \quad j = 1, 2, 3, \dots, L. \quad (6)$$

Combining the fuzziness measurement formula with the confidence level, we can redefine the classification fuzziness measurement formula. See the following formula: for details:

$$G(E) = g(\pi(C_t|E)). \quad (7)$$

At this time, assuming that the value of variable A on set u is F and the fuzzy semantic item set corresponding to variable B is p , the calculation formula of fuzzy classification and division of the correlation between variables A and B is

the fuzzy decision tree algorithm rules can be seen in the following formula:

$$G(P|F) = \sum_{t=1}^k w(B_t|F)G(B_t|F), \quad (8)$$

where $w(B_t|F)$ represents the size of fuzzy set F , see formula (9) for specific calculation method, and see formula (10) for fuzzy evidence at this time:

$$w(B_t|F) = \frac{M(B_t \cap F)}{\sum_{j=1}^k M(B_j \cap F)}, \quad (9)$$

$$\mu_{E_a}(u) = \begin{cases} \mu_E(u), & \mu_E(u) \geq \alpha, \\ 0, & \mu_E(u) < \alpha. \end{cases} \quad (10)$$

Finally, we bring the preprocessed data into the semantic item membership function to complete the data fuzzification, which obtains the parameters by using the Kohonen feature mapping algorithm and realizes the transformation from the fuzzy decision tree algorithm to fuzzy rules. The induction process of a fuzzy decision tree consists of the following steps: (1) data preprocessing; (2) induction and establishment of decision tree; (3) The obtained fuzzy decision tree is transformed into a set of fuzzy rules; and (4) The obtained fuzzy rules are applied to classification. The membership function of semantic items is shown in the following formulas (11)–(13):

$$\mu_{T_k}(x) = \begin{cases} 1, & x \geq m_k, \\ \frac{(x - m_{k-1})}{(m_k - m_{k-1})}, & m_{k-1} < x < m_k, \\ 0, & x \leq m_{k-1}, \end{cases} \quad (11)$$

$$\mu_{T_1}(x) = \begin{cases} 1, & x \leq m_1 \\ \frac{(m_2 - x)}{(m_2 - m_1)}, & m_1 < x < m_2 \\ 0, & x \geq m_2, \end{cases} \quad (12)$$

$$\mu_{T_i}(x) = \begin{cases} \frac{(m_{i+1} - x)}{(m_{i+1} - m_i)}, & m_i < x < m_{i+1}, \\ \frac{(x - m_{i-1})}{(m_i - m_{i-1})}, & m_{i-1} < x < m_i, \\ 0, & x \geq m_{i+1} \cup x \leq m_{i-1}. \end{cases} \quad (13)$$

In order to verify the member experience quality evaluation of the MOOC teaching model of basic education based on fuzzy decision tree algorithm, this study designs the following experiments to evaluate the model. Firstly, the data mining classification accuracy of fuzzy decision tree is calculated by the tenfold cross validation method. In order to prove the accuracy of the fuzzy decision tree algorithm, in addition to using the fuzzy decision tree algorithm to process data, we also use the

decision tree algorithm to fully prove the accuracy of the algorithm. We randomly selected three different data sets, session V , session M , and session A . Where session V represents all access records under a video, session M represents access records of different device types, and session A represents the video access records in different regions. By changing the parameters, observe the changing trend and fitting accuracy of the fuzzy decision tree model on the three different data sets of session V , session M , and session A , and compare the dependent variables under different conditions, so as to judge the impact of the quality of the video, the region of the members, and the equipment of the members on the accuracy of the prediction results of the model Figure 3.

After obtaining the preliminary verification results, we need to consider its reliability. Therefore, we continue to compare at different levels: significance level a and confidence level B . When a and B take different values respectively, the prediction accuracy of fuzzy decision tree models session a , session m , and session V changes. The results are shown in Figure 4. Then, taking conversation a as an example, we show different significance levels a and confidence levels B . When the accuracy changes, we can see that whether the value of B is any of 0.2 to 0.8, the session set prediction accuracy remains at about 43% without obvious fluctuations. When the value of a is 0.5, it is different from B . when the value of a is 0.8, the prediction accuracy of the session begins to differ at the level of B . The maximum prediction accuracy of session a is 44%. However, compared with Class A and class B , the prediction accuracy under Class A and class B still has no significant difference. When the value of a is 0.6, the prediction accuracy of level B for level a sessions is higher. However, when the b value is 0.8, the prediction accuracy of level a sessions is lower than that of level B sessions. The prediction accuracy of level a sessions is 0.2-0.6. When the value of a is 0.9 and the value of B is 0.2, the prediction accuracy of session a is the highest, which is 67%. When e of a is 0.1 and B is 0.2, the prediction accuracy is the lowest, which is 42%. The predicted change law of session a is basically the same as that of session C .

4. Result Analysis and Discussion

In the MOOC teaching model of inquiry basic education, in the experience quality evaluation system of members for online video teaching, if only the technical needs of data mining are considered, the decision tree algorithm can be met in the commonly used data mining technology. However, because the decision tree algorithm is only used for the processing of discrete attribute data sources in practice, it has some limitations on the processing of continuous variables, such as access time. Therefore, based on the existing data mining technology, this paper proposes a fuzzy decision tree algorithm. The specific advantages and disadvantages of the algorithm and the decision tree algorithm have been compared in detail at the beginning of the third part, so it is not repeated here. Only the comparison results of the prediction accuracy of the decision tree algorithm and the fuzzy decision tree algorithm on the three sets of session A , session M , and session T are presented. The specific results are shown in Figure 5. From the figure, we can see that for different

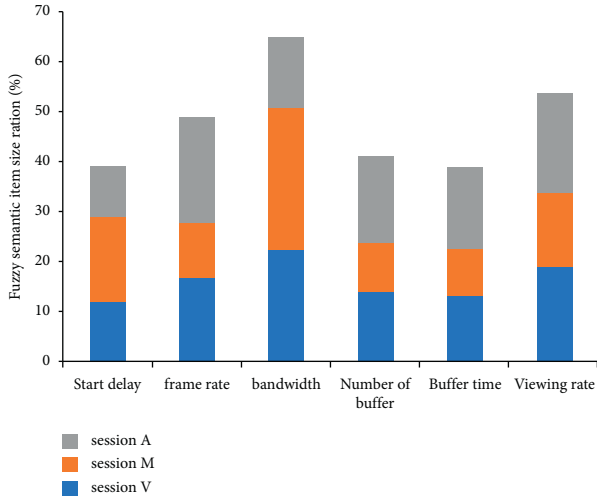


FIGURE 3: Different sets of video viewing index size proportional relationship.

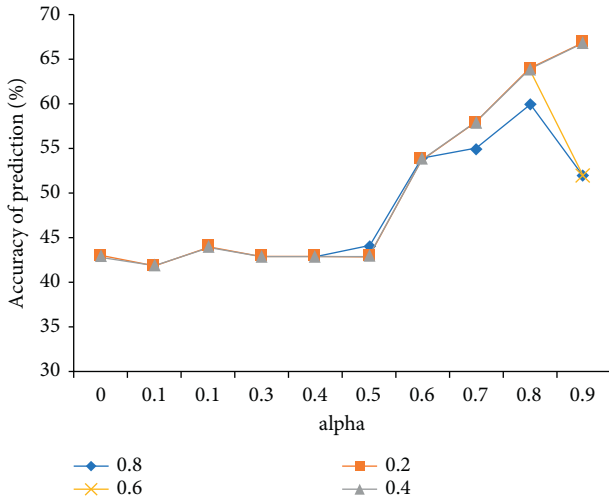


FIGURE 4: Fuzzy decision tree model predicts accuracy when parameters are different.

data sets session *A*, session *M*, and session *T*, the model prediction accuracy based on fuzzy decision tree algorithm is always higher than that based on decision tree algorithm, which further verifies the rationality of our algorithm. In addition, among different sets, the model prediction accuracy of session *M* based on the fuzzy decision tree algorithm is the highest, reaching about 81%, while the model prediction accuracy of session *T* based on the fuzzy decision tree algorithm is the lowest, about 63%. It is preliminarily speculated that this is due to the different degree of analysis of the video quality capture of members' access and the capture accuracy of members' access to geographical locations. This problem can be further explored in subsequent research.

In addition, we notice that the classification accuracy of fuzzy decision tree is also different on different subsets and complete sets. It can be found from Figure 6 that the prediction accuracy of the accuracy prediction model on subsets session *A*, session *M*, and session *T* far exceeds that

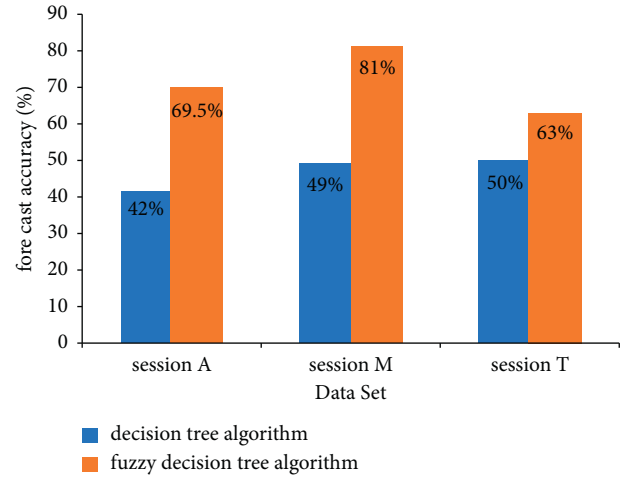


FIGURE 5: Comparison of prediction accuracy between fuzzy decision tree and decision tree model.

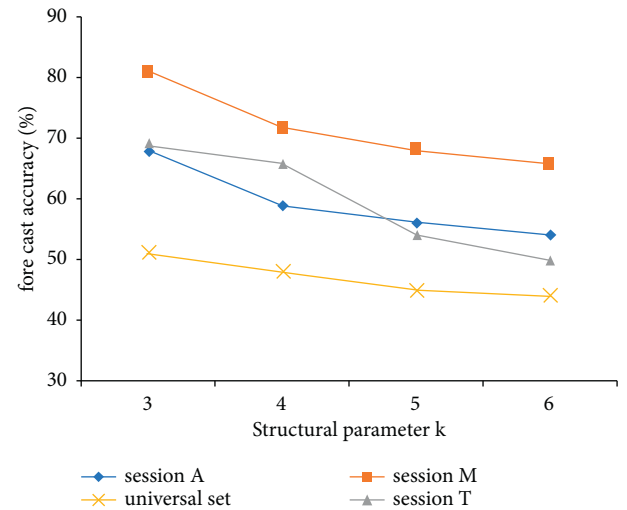


FIGURE 6: Comparison of model prediction accuracy under different subsets and different K values.

on the full set *S*. This result shows that the experience of members accessing the online video teaching platform is greatly related to the video content, region, and equipment accessed by members. The gap between each factor is large, and there is interaction. We speculate that the reasons are as follows: Firstly, it is related to the members' own interests. The interest in video affects the members' tolerance of video. Secondly, it will be affected by the background, cultural habits, and local policies of the member's location. For example, there are great differences in the learning methods between the first tier cities and the second and third tier cities. Due to development constraints, the second and third tier cities may not understand the online platform channels, so the visit will also be affected. Or if the local requires students to teach on average, the local visit volume will also be greatly affected. In addition, using different devices to access the MOOC platform also represents the current state of

members to a certain extent, so it will affect the accuracy of the evaluation model. Starting from practical reasons and exploring from the perspective of fine granularity, we can still find out other objective influencing factors. Due to the length of this paper, we will not explore in detail.

5. Conclusion

The main goal of this paper is to establish a set of evaluation teaching models suitable for the development of China's industry and combined with China's national conditions. Considering the data source, the accuracy of the results, and other factors, we chose the industry head platform MOOC network as an example. Through the exploration of the MOOC teaching model of basic education and based on the fuzzy decision tree algorithm, we establish the member experience evaluation model. Finally, experiments show that the accuracy of the model is good and the reliability of the prediction results is high. However, there are many online Internet platforms in China, and with the further development of online video teaching industry, the characteristics of each platform are also different. Therefore, whether the fuzzy decision tree model proposed in this paper can be applied to other online basic education models remains to be discussed. The research scope can be further expanded in future work.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] J. Li, "Application of intelligent fuzzy decision tree algorithm in English teaching model improvement," *Complexity*, vol. 2021, Article ID 8631019, 69 pages, 2021.
- [2] Z. Gu and C. He, "Application of fuzzy decision tree algorithm based on mobile computing in sports fitness member management," *Wireless Communications and Mobile Computing*, vol. 2021, no. 22, 60 pages, Article ID 4632722, 2021.
- [3] M. Farnaz, M. Maryam, R. S. Meysam, and P. WitoldChimFlexDT, "Chi-square-based multi flexible fuzzy decision tree for data stream classification," *Applied Soft Computing Journal*, vol. 105, no. 7, Article ID 107301, 2021.
- [4] G. Hu, S. Mohammadiun, A. A. Gharahbagh, J. Li, K. Hewage, and R. Sadiq, "Selection of oil spill response method in Arctic offshore waters: a fuzzy decision tree based framework," *Marine Pollution Bulletin*, vol. 161, no. 19, Article ID 111705, 2020.
- [5] H. Aggarwa, H. D. Arora, and V. Kumar, "An entropy based decision making problem using fuzzy decision tree classification of data mining," *Journal of Critical Reviews*, vol. 7, p. 18, 2020.
- [6] A. Vijaya Lakshmi and V. S. Ghali, "Fuzzy decision tree based characterization of subsurface anomalies," *International Journal of Recent Technology and Engineering*, vol. 8, pp. 5–7, 2020.
- [7] J. Sun, "Power instability prediction method for wind turbine based on fuzzy decision tree," *Journal of Intelligent and Fuzzy Systems*, vol. 39, pp. 2–5, 2020.
- [8] D. Teekaraman, S. Sendhilkumar, and G. S. Mahalakshmi, "Semantic provenance based trustworthy users classification on book-based social network using fuzzy decision tree," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 39, no. 28, pp. 41–44, 2020.
- [9] J. Shanmugam and V. S. Ghali, "Fuzzy decision tree based characterization of subsurface anomalies," *International Journal of Recent Technology and Engineering*, vol. 8, pp. 51–60, 2020.
- [10] G. Fison and P. Paci, "SAveRUNNER: an R-based tool for drug repurposing," *BMC Bioinformatics*, vol. 22, no. 1, pp. 1–10, 2021.
- [11] R. Jan, V. Levashenko, E. Zaitseva, M. Kvassay, and S. Subbotin, "Non-destructive diagnostic of aircraft engine blades by fuzzy decision tree," *Engineering Structures*, vol. 197, no. 19, pp. 77–79, 2019.
- [12] Y. Mu, X. Liu, L. Wang, and A. B. Asghar, "A parallel tree node splitting criterion for fuzzy decision trees," *Concurrency and Computation: Practice and Experience*, vol. 31, pp. 17–20, 2019.
- [13] Z. Shi and C. Huang, "Network video quality assessment method using fuzzy decision tree," *IET Communications*, vol. 13, pp. 14–16, 2019.
- [14] N. Khazali, M. Sharifi, and M. Ali Ahmadi, "Application of fuzzy decision tree in EOR screening assessment," *Journal of Petroleum Science and Engineering*, vol. 177, pp. 7–12, 2019.
- [15] H. Zheng, H. Jing, Y. Zhang, G. Huang, Z. Zhang, and Q. Liu, "A general model for fuzzy decision tree and fuzzy random forest," *Computational Intelligence*, vol. 35, pp. 2–5, 2019.

Research Article

Generative Adversarial Network for Musical Notation Recognition during Music Teaching

Na Li 

School of Music and Performing Arts, Mianyang Teachers' College, Sichuan, Mianyang 621000, China

Correspondence should be addressed to Na Li; lina@mtc.edu.cn

Received 16 April 2022; Accepted 14 May 2022; Published 7 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Na 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 order to improve the quality and efficiency of music teaching, we try to automate the teaching of music notation. With the addition of computer vision technology and note recognition algorithms, we improve the generative adversarial network to enhance the recognition accuracy and efficiency of music short scores. We adopt an embedded matching structure based on adversarial neural networks, starting from generators and discriminators, respectively, to unify generators and discriminators from the note input side. Each network layer is then laid out according to a cascade structure to preserve the different layers of note features in each convolutional layer. Residual blocks are then inserted in some network layers to break the symmetry of the network structure and enhance the ability of the adversarial network to acquire note features. To verify the efficiency of our method, we select monophonic spectrum, polyphonic spectrum, and miscellaneous spectrum datasets for validation. The experimental results demonstrate that our method has the best recognition accuracy in the monophonic spectrum and the miscellaneous spectrum, which is better than the machine learning method. In the recognition efficiency of note detail information, our method is more efficient in recognition and outperforms other deep learning methods.

1. Introduction

Computer simulations play a very important role in teaching music today. The greatest advantage of multimedia technology is cross-media presentation. The traditional school board is limited to written and face-to-face instruction, and much virtual knowledge cannot be displayed, so students cannot feel the main points of learning in a personal way. Thus, the development of computer technology today has made education presentable, and it can realize the conversion from text data to image video, also the conversion of music notation to audio, and also the virtual performance of musical instruments. The combination of computer technology and education has added more fun to education, and the novel teaching methods can mobilize students in all aspects, make the teaching mode more active, and improve the quality of teaching significantly.

The traditional model of music education is one-on-one teaching between teacher and student, and music-type art training is usually a small course. Lessons involving music

notation are rather boring, and classroom-style instruction prevents students from getting a first-hand feel for the tones and distinctions of each musical notation. All music notation and musical notation require rote memorization by students to remember, which drastically reduces the quality of music notation instruction [1]. The advent of computer technology has improved the efficiency of teaching music notation [2]. With the aid of computer technology, various music software was born to bring great convenience to students' extracurricular music learning. Music software contains virtual instrument functions, digital score presentation functions, virtual arranging tools, and digital tuners. A series of music assistance programs have emerged to make self-learning music more convenient and are sought after by a large number of amateur music lovers [3].

Music notation initially existed in the form of handwriting, the construction of music scores in the early period of music festivals was based on the sense of sound, and the writing of all music symbols varied from person to person, which brought great challenges to the work of automatic

music notation recognition. The fusion technology of computer vision and image recognition algorithm to realize the music score recognition can improve the working efficiency and save the working cost. To orderly carry out the music score recognition work and solve the problem of variation of handwritten scores, the researcher specially designed a notation frame, and each music notation frame can learn the handwriting habits of different musicians independently and will automatically label the information of music source, author, and genre after recognition [4–6]. The computer input work is performed by scanning the music notation manuscript, and each music symbol is designed in advance as a label for easy learning of the score. After the recognition is completed, musicians can download and edit the music from the Music Resource Sharing website.

Music notation recognition systems give more prominence to image recognition techniques, yet ignore the homogeneous heterogeneity of music notation handwriting. For this problem, the researchers used computer vision techniques for the same specification. Each handwritten music notation has a different representation, and different types of music notation cannot be processed with the same heuristic [7]. According to the frequency of music notation usage, the scanned music notation needs to be reconstructed phonetically, and the music notation is assigned according to different weights as a way to cater to the fluency of the music notation. With the development of machine learning techniques, each note symbol in a musical notation can be generalized and model training can be achieved by manually building a note symbol dataset. So far, some researchers have demonstrated that the combination of note symbols and machine learning techniques is not suitable for all music notation, where note symbols captured by pentatonic scores need to be preprocessed with images to be applicable, and special music element classification requires dataset-independent production based on specific notes. Although the machine learning method is the first innovation in the research of music short score recognition, the method adds a lot of work cost to the method due to the lack of a clear and explicit engineering framework and the tedious manual labeling work, plus the accuracy and real-time performance of the machine learning method are not good enough. Therefore, there is still a lot of research space in the field of music score recognition [8, 9].

In the experiments of adversarial generative networks, we try to fuse music notation features into the generator to accomplish the task of music notation recognition by stimulating the generation of pseudo-samples. The musical notation recognition method we designed consists of symbol recognition and score reconstruction. The flowchart is shown in Figure 1. We adopt an embedded matching structure based on adversarial neural networks, starting from generators and discriminators, respectively, to unify generators and discriminators from the note input side. Then, each layer of the network is laid out according to the cascade structure, and then the residual blocks are inserted in some network layers to break the symmetry of the network structure and enhance the ability of the adversarial network to acquire note features. Finally, we validate the

effectiveness of our method on a public dataset of music notation.

The rest of the paper is organized as follows. Section 2 introduces the research history and research results of musical notation recognition. Section 3 details the principles and implementation procedures related to the improved adversarial music notation recognition network. Section 4 shows the experimental datasets and the analysis of the experimental results. Finally, Section 5 summarizes our research and reveals some further research work.

2. Related Work

There are many branches of music notation recognition research, among which optical music recognition is one of the well-known research branches. Optical music recognition mainly relies on optical factors to achieve the recognition of music symbols. The literature [10] has a very in-depth study on optical music recognition, and the authors try to establish a series of different stages to deal with the grading of music symbols. Considering the differences between low-order to high-order notes, the authors propose a general optical music recognition framework and supplement it with different note segmentation methods to achieve the task of nondifferentiated recognition of musical short scores. For special notes, the authors default to note segmentation for preprocessing and then unify all music symbols and treat them as independent phonemes for optical scanning.

Researchers in literature [11] tried to improve the recognition accuracy of music notation from the perspective of images, and they proposed a binarization algorithm, which can temporarily solve the problem of a small number of music notation images. However, the method lacks generalization, has limited applications, requires adaptation for note images with different characteristics, relies heavily on optical music recognition methods in the conversion between high and low order for different music scores, and becomes less applicable due to the lack of flexibility of a unified note recognition framework. Researchers in the literature [12–14], after validating a large number of machine learning methods, found that DNN models have a high degree of generalizability and are better at musical score recognition with better recognition accuracy. The application potential of deep learning methods in music short scores is high, but the method requires more stringent datasets. In the construction of the musical score dataset, the skewed manuscript content needs to be corrected in advance, and the overlapping notes need to be separated in advance according to the correct score. In addition, to improve the inclusiveness and increase the volume of the dataset, the literature [15, 16] proposed data enhancement algorithms to improve the coverage of different angles and sizes of note features in the dataset.

The study of musical notation is not a smooth process, in which music notation segmentation is a great challenge. Music notation is different from characters, which have a professional character library that can be used as a database to unfold the mapping. However, music symbols are a new

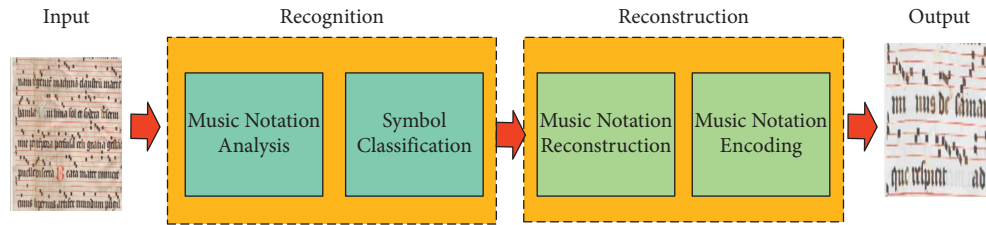


FIGURE 1: Music notation recognition process.

task in the early stage of research, and without a huge music symbol library as data support, the music symbol database needs to be built from scratch, which undoubtedly brings great difficulties to the work of music score recognition. At the initial stage of music symbol database establishment, researchers chose to define the scanned documents of notes with poor image quality with small element labels and then separated the notes from pseudo-notes by noise threshold. Based on the former research method, the literature [17] proposed the principle of object reconstruction at the initial stage of notes, which solved the problem that the overlapping of notes and pseudo-notes at the initial stage of notes could not be separated. Considering that the workflow of the traditional note detection method is too tedious and the accuracy is low, the literature [18, 19] first tried to apply the region-based neural network method to note feature extraction, abandoning the note separation step, directly starting from note features, and training the model to directly complete note recognition. However, the method has high requirements on the dataset. Researchers in the literature [20] were again inspired by the musical notation baseline and used the baseline as the note criterion to build a variety of note and notation models to achieve excellent note recognition accuracy with an adaptive fusion approach.

To avoid differentiated conversions between higher and lower orders of notes, researchers in the literature [21, 22] chose a neural network approach that starts with the overall musical notation. They transformed the output of the neural network as a sequence of notes and then annotated the notes of higher and lower orders in the sequence position. The released note elements are converted to actual notes in backpropagation, preventing multiple repetitions of detection during postprocessing. The experimental results demonstrate that the method is faster and takes less time to detect in music notation detection. To ensure the wholeness of music notation in note separation, the literature [23] proposed a method to reshape music notation using hidden Markov models and achieved good results in experiments. The researchers in the literature [24] transferred the method to the handwritten data of music notation based on the former and proposed a fusion algorithm of neural network and discriminative algorithm [25], which was able to identify the handwritten data completely and accurately under the ideal situation of unstructured environmental factors.

Among the methods of music notation recognition, most researchers prefer the end-to-end neural network method because the music notation needs to do note segmentation process during the preprocessing process, due to the variability problem of manuscript notes. In the process of note image

acquisition, poor image quality, note overlap, note occlusion, and other problems can affect the integrity of note segmentation work. To solve this problem, researchers have used end-to-end neural network recognition methods to extract local and overall note features to ensure the integrity of local features and correct recognition of damaged notes [26, 27]. In addition, researchers in the literature [28, 29] proposed a deep neural network-based note synthesis method for the problem of damaged notes that cannot be correctly recognized, which is mainly based on the local features of notes and automatically improves the defective notes against the note learning library, which has a high dependency on the note learning library. In other words, the richness of the note learning library determines the note recognition accuracy of the method.

3. Method

3.1. Basic Pipeline. To ensure the feature integrity of the musical score, a generative adversarial network is chosen as the basis for learning from note local features through an unsupervised training mode, which can perform undifferentiated feature reorganization for various images with the aid of computer vision techniques. The generative adversarial network consists of two parts, a generator and a discriminator, which are used to simulate the note features to be learned and generate pseudo-samples with high feature similarity to match them. In the data input stage, only the real note samples that have been preprocessed are segmented, and then the generator simulates similar pseudo-samples based on the feature analysis. The discriminator will score the similarity between the fake samples and the real samples, and the fake samples that meet the specified scoring range will be output directly through the simulator, while the fake samples that do not meet the scoring range will be fed back to the front end to regenerate the fake samples until the fake samples that meet the scoring criteria are generated. The structure of the generative adversarial network is shown in Figure 2.

Generative adversarial networks are based around mutual game learning methods as mathematical principles and are effective in obtaining independent note features in music notation recognition work. For note separation of pentatonic and characteristic scores, the generative adversarial network will optimize the feature editing of the characteristic notes using a pseudo-sample generation model to control the sample output of the note features by editing between discriminator parameters. Such an approach can indirectly control the local and overall features from the notes, avoiding the problem of feature loss in feature separation.

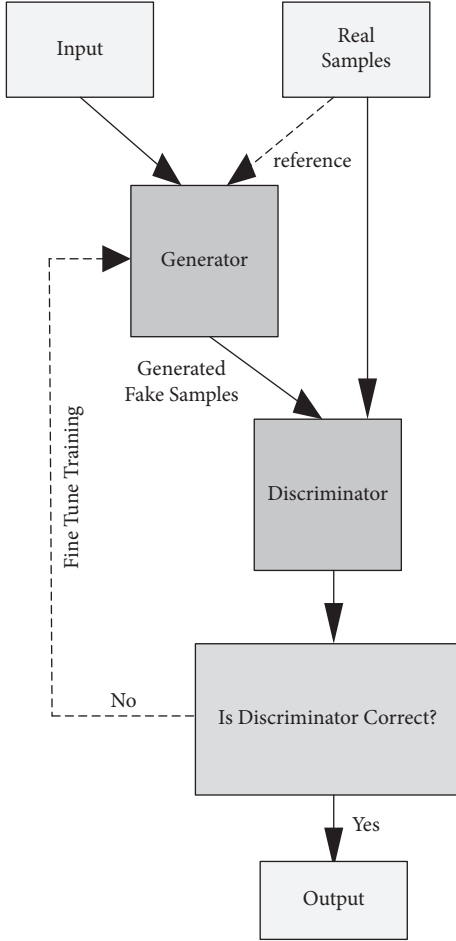


FIGURE 2: Generative adversarial network architecture.

3.2. Generator. The generator is a combination of a fully connected neural network and a deconvolutional network. The generator can automatically obtain the key features of the input notes and spectral data based on them and generate pseudo-samples with similar features at the terminals. The principle of generator action is shown in Figure 3. For the pseudo-sample output by the generator, we will discriminate the feature parameters from different dimensions and feedback to the training layer to adjust the feature dimension parameters to get better quality pseudo-samples.

Researchers in the literature [30, 31] aimed to implement the embedded matching problem in feature encoding and decoding. They designed feature encoders with similar specifications at the input and consistent feature decoders at the output, and experimentally demonstrated that such a matching design improved the efficiency of pseudo-sample generation in generative adversarial networks, reduced the number of parameters, and improved the robustness of the networks. Therefore, we also adopt the same combination of decoding and encoding embedded matching, and to make the note feature encoding more adaptable to the embedded model, we also adopt the cascade structure as the network skeleton connection. The input of the encoder is an independent downsampled convolutional layer that can retain the expressed intensity features of the input note features,

assuming that the retained features are I^{low} . After down-sampling, the note features will be converted to the hidden layer as a backup. The literature [32] refers to the residual network in the structure design of the decoder, which avoids the problem of note feature information omission during the decoding training process and makes the whole decoding network more compact. We also adopt the same decoder design strategy, and we introduce different levels of residual blocks in the decoder to ensure that note features of different strengths can be fully decoded. In addition to the introduction of residual blocks, the decoder uses the upsampling deconvolutional layer as the main network to realize the conversion between the prescribed expression ranges of note features of different intensities. The convolutional layers in the decoder and encoder uniformly employ normalization operations and ReLU linear activation with a step size of 2. We used the X conv operator mentioned in the literature [33], assuming that the given K input is (p_1, p_2, \dots, p_k) and the K input is the result of a multi-layer perceptron weighting. Then, the $K \times K$ transform matrix $X = \text{MLP}(p_1, p_2, \dots, p_K)$ is executed and the convolution summation gives the transformed features of the convolution operator X . To solve the adjacency effect between different note features, we have the following mathematical definition for the X conv operator.

$$F_p = X_{\text{conv}}(K, p, P, F),$$

$$X_{\text{Conv}}(K, p, P, F) = \text{Conv}(K, \text{MLP}(P - p) \times [\text{MLP}_\delta(P - p), F]), \quad (1)$$

where p denotes the note feature points, K denotes the adaptive convolution kernel, $P = (p_1, p_2, \dots, p_k)^T$ denotes the K points in its neighborhood, and $F = (f_1, f_2, \dots, f_K)$ denotes the features of different notes. Using the principle of X conv operator, we construct a musical short note generator representing different intensity features, as shown in Table 1. We replace the connection of the encoder and decoder and use a jump connection structure to ensure that the location information of the random note features matches each other.

3.3. Discriminator. The discriminator is the same as the generator and has the same deconvolution network hierarchy. The discriminator evaluates the pseudo-sample output by the generator by using the feature parameters of the real samples as the discriminant criteria. If the evaluation result is not up to the standard, the pseudo-sample is fed back to the generator and the pseudo-sample is generated again. The discriminator is capable of adjusting the parameters according to the note characteristics on its own or manually on demand. The working principle of the discriminator is shown in Figure 4.

The conversion between low-order note features and high-order note features is prone to pitch confusion, and we filter the high-order note features in the generator, compensate the high-order note features by the underlying data density, and distinguish the similarity between high-order and low-order note features by the high-density note feature layer in the discriminator. Researchers in the literature [34]

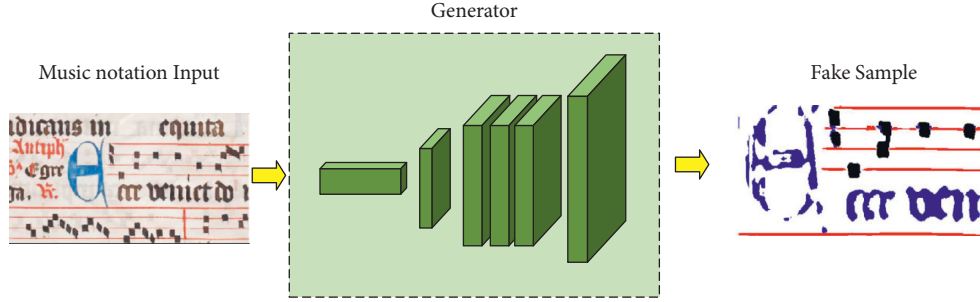


FIGURE 3: Music notation generator process.

TABLE 1: Detailed hierarchy of generators.

Layer	Type	Detail
1	Input	5001 × 3
2	X conv	Np = 2400, C = 16, K = 8, D = 1
3	X conv	Np = 800, C = 64, K = 16, D = 2
4	X conv	Np = 200, C = 256, K = 24, D = 2
5	X conv	Np = 800, C = 64, K = 24, D = 2
6	X conv	Np = 2400, C = 16, K = 16, D = 2
7	X conv	Np = 5001, C = 16, K = 16, D = 1
8	Fully connected	C = 3
9	Output	5001 × 3

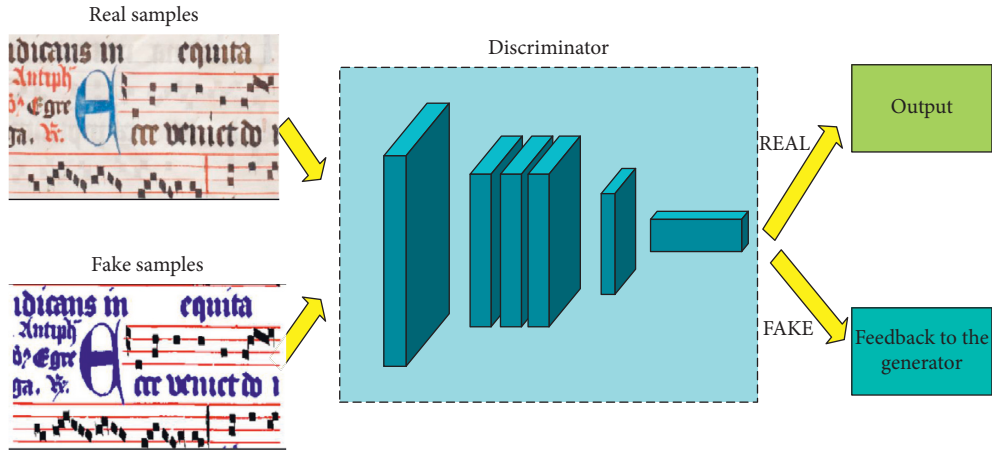


FIGURE 4: Music notation discriminator process.

proposed an alternating training model on the problem of optimizing note feature discretization and replaced the mathematical computation in principle with the iteration of maximum and minimum values. We have adopted the same approach, and we have the following mathematical definition for the maximum and minimum value turnover in the evaluation of high- and low-order note features.

$$\frac{\min}{\text{Gen}} \frac{\max}{\text{Dis}} = E_I^{\text{high}} \log(\text{Dis}(I^{\text{high}})) + E_I^{\text{low}} \log(1 - \text{Dis}(\text{Gen}(I^{\text{low}}))), \quad (2)$$

where Gen denotes the note features generated by the generator and Dis denotes the note features determined by the discriminator. $\{I^{\text{low}}, I^{\text{high}}\}$ denotes a pair of musical notations with different feature strengths but the same note

order. The adversarial loss function equations for generator Gen and discriminator Dis are shown below.

$$L_{G,\text{adv}} = -\frac{1}{N} \sum_{n=1}^N \log(\text{Dis}(\text{Gen}(I_n^{\text{low}}))),$$

$$L_{D,\text{adv}} = -\frac{1}{N} \sum_{n=1}^N \{\log(\text{Dis}(I_n^{\text{high}})) + \log(1 - \text{Dis}(\text{Gen}(I_n^{\text{low}})))\}, \quad (3)$$

where N denotes the total number of training note samples. In the process of adversarial neural network convergence, different discriminator parameters are set according to different note strata, and hierarchical restriction means are adopted for pseudo-sample convergence to screen high-

order notes and feedback to the generator to generate high-density note features. For this purpose, we established discriminator network layers with different hierarchical structures, and the network layer density information is shown in Table 2.

3.4. Loss Function. There is a clear problem of differential differentiation between high-order and low-order notes in the hierarchical feature representation, and the real note feature I^{high} , modulated by the high-intensity density parameter, guides the generator to synthesize pseudo-samples with highly similar intensity of feature $\text{Gen}(I^{\text{low}})$. The literature [35] mentions a point-by-point loss optimization approach in the pseudo-sample optimization strategy, which constrains the loss function by controlling the relative distance between high-intensity features and low-intensity features. In this paper, we control the feature distance between high-intensity and low-intensity features, constrain the features using the L1 loss function, and increase the integration of features of different classes using the L2 loss function. Our loss function constraint equation is shown below.

$$L_{\text{note}} = \frac{1}{N_{\text{note}}} \sum_{i=1}^{N_{\text{note}}} \left\| \text{Gen}(I^{\text{low}})_i - I_i^{\text{high}} \right\|, \quad (4)$$

where N_{note} denotes the note features in the low-order samples and also denotes the tone spectrum data points in the high-order samples. Combining the above loss functions, the systematic loss function formula of our optimized generative adversarial network is as follows.

$$L = \omega_1 L_{G\text{-adv}} + \omega_2 L_{\text{note}}, \quad (5)$$

where ω_1 denotes the weighting coefficient. We adopt the alternating training network iteration mode, and the generator-side network can generate pseudo-samples with very high feature similarity in iterations, which can reduce the discriminator parameter adjustment step when discriminating with the real samples.

3.5. Music Notation Recognition Network. For applying a deep neural network model to the recognition of short scores for music teaching, we compared several neural networks in the selection of the underlying network and finally chose a generative adversarial network. The most unique advantage of the generative adversarial network is that it does not affect the original note feature structure, which is regenerated by a generator simulating real samples. We propose an improved generative adversarial network method based on this network to improve the recognition accuracy and recognition speed of music notation. In our improved strategy, the generator and discriminator are embedded together in the residual structure, which can successfully resolve the recognition differences between notes with different data densities. For low-order note data, the convolutional neural network can generate auxiliary samples by downsampling. For high-order note data, the inverse convolutional network upsampling can get the note feature intensity, and then the

TABLE 2: Detailed hierarchy of discriminator.

Layer	Type	Detail
1	Input	5001×3
2	X conv	$N_p = 2400, C = 16, K = 8, D = 1$
3	X conv	$N_p = 800, C = 64, K = 16, D = 2$
4	X conv	$N_p = 200, C = 256, K = 24, D = 2$
5	Fully connected	$C = 1$
6	Mean	—
7	Output	1×1

pseudo-samples can be generated by the feature calculation through the X conv operator. In the joint output, different layers of music notation are modeled and filtered with features in the form of note features, and the classifier obtains key features from real samples to provide guidelines for pseudo-sample generation, fusing comprehensive note features. The detailed music notation recognition network is shown in Figure 5.

4. Experiment

4.1. Datasets. To validate our method for music notation score recognition, we chose a public dataset for experimental validation. The dataset of the music notation series contains 4 categories, which are a monophonic spectrum, polyphonic spectrum, polyphonic spectrum, and mixed spectrum. The most representative dataset in the monophonic category is the Bach Chorales (BC) dataset [36], which is in XML format for the whole series and contains four vocal parts and multiple melodic parts. This dataset has an important role in the melodic generation and harmonic modeling studies. The most famous dataset for polyphonic spectra is the MAE-STRO (MO) dataset [37], which is a collection of MIDI-enabled piano melodies, each corresponding to a different audio spectrum, and on which many of Google’s spectral studies have been conducted. The most representative dataset for polyphonic scores is the Video Game (VG) dataset [38], where most of the scores are derived from video game music and are mainly used for electroacoustic synthesis. The most famous dataset for mixed scores is the Lakh (LH) dataset [39], which has the advantage of a large number and is mostly used for model pretraining. Besides, we added a wild dataset the Largest MIDI (LM) dataset [40] to ensure the diversity of music notation and to improve the generalization of the music notation recognition model. Details are dataset information as shown in Table 3.

4.2. Analysis of Results. To verify the effectiveness of our method for note recognition in music notation, we compared machine learning methods and deep learning methods. Among the machine learning methods, we chose the most representative logistic regression (LR) and decision tree (DT), and among the deep learning algorithms, we chose recurrent neural network (RNN) and long short-term memory network (LSTM). To ensure independent validation relationships between each method, we conducted five sets of experiments during the training process to independently

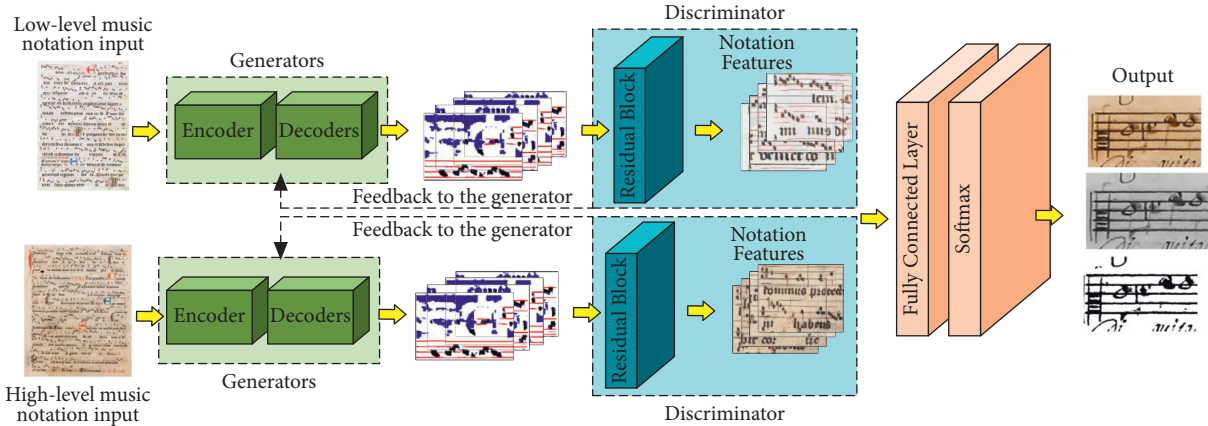


FIGURE 5: Music notation recognition network.

TABLE 3: Dataset information.

	Datasets				
	BC	MO	VG	LH	LM
Train	79801	56342	46351	87500	65492
Test	21420	15993	20365	30021	29564
Total	101221	72335	66716	117521	95056

verify the efficiency of each group of methods for sound spectrum recognition. We use recognition accuracy (P), F1 score, and recall rate (R) as the evaluation criteria of the music short score recognition methods. Each method detection result will be directly fed into the statistical calculation part of the dataset, and the final evaluation result will be obtained by the balance between the total number and quality of the dataset. To verify the preference of each method in different source datasets, we divided the dataset into two groups. The first group is composed of monophonic scores, polyphonic scores, and polyphonic scores, and this dataset is mainly used to verify the efficiency of the music notation recognition methods for independent recognition of monophonic and polyphonic notes. The experimental results are shown in Table 4.

From the experimental results in the above table, it can be seen that the machine learning method does not perform well enough in the independent recognition experiments for both monophonic and polyphonic spectra, and the accuracy is below 70%. The deep learning method maintains the recognition accuracy between 70% and 86% in the spectrum recognition experiments, and our method achieves an average recognition accuracy of 90% in the spectrum recognition. The experimental results demonstrate that our method has the best independent recognition in both monophonic and polyphonic phonetic spectra. In the second experimental dataset, we chose mixed and wild tone spectra as the base dataset, and this set of experiments is mainly to verify the recognition effect of the tone spectral recognition method in the miscellaneous tone spectra. The experimental results are shown in Table 5.

From the experimental results in the table above, it can be seen that the machine learning method is less efficient in the recognition of the murmur spectrum than the

monophonic and polyphonic spectra, and the deep learning method performs generally in the experimental results of the murmur spectrum, with the overall average recognition accuracy remaining at 78%, while our method performs even better in the recognition of the murmur spectrum, with the overall recognition accuracy remaining above 90%. Since our method adopts the separated feature twin method, which does not affect the original note features, it is more efficient in the process of murmur note feature extraction and has higher recognition accuracy. The combined results of all experiments show that our method is better compared to both machine learning methods and deep learning methods.

To verify whether the note information of music notation is accurately recognized, we selected four metrics from the note level: note meta information (NMI), note nodal line (NNL), note chord (NC), and note segmentation (NS). In order not to let the difference in datasets affect the efficiency of note recognition for each method, we selected a common dataset from the monophonic group and the murmuring group for validation, respectively. After the previous experiments, we found that there are significant differences between machine learning methods and deep learning methods. To save experimental costs, this session of experiments will only validate the note recognition efficiency of deep learning methods. The experimental results are shown in Table 6.

The experimental results in the above table show that the overall recognition rate of note details of monophonic notes is higher than that of the miscellaneous note spectrum. The reason for this result is that the monophonic score is more standardized in the segmentation of note detail information, while the miscellaneous score is a mixed scale, which is not standardized in the segmentation, causing the problem of

TABLE 4: Comparison of single-tone and multi-tone spectra.

	BC			MO			VG		
	P	R	F1	P	R	F1	P	R	F1
LR	0.57	0.58	0.64	0.63	0.67	0.64	0.63	0.58	0.61
DT	0.69	0.70	0.69	0.70	0.61	0.64	0.68	0.64	0.64
RNN	0.75	0.81	0.82	0.76	0.70	0.71	0.84	0.81	0.77
LSTM	0.81	0.85	0.84	0.81	0.79	0.81	0.86	0.83	0.83
Ours	0.92	0.93	0.88	0.88	0.89	0.87	0.91	0.93	0.91

TABLE 5: Comparison of the recognition effect of the miscellaneous sound spectrum.

	LH			LM		
	P	R	F1	P	R	F1
LR	0.51	0.53	0.61	0.62	0.53	0.54
DT	0.60	0.63	0.64	0.69	0.62	0.59
RNN	0.71	0.80	0.81	0.77	0.75	0.73
LSTM	0.79	0.83	0.84	0.82	0.82	0.75
Ours	0.90	0.95	0.93	0.92	0.90	0.95

TABLE 6: Notation recognition efficiency of different methods.

	NMI		NNL		NC		NS	
	BC	LH	BC	LH	BC	LH	BC	LH
RNN	0.61	0.51	0.71	0.68	0.81	0.75	0.53	0.41
LSTM	0.73	0.64	0.82	0.75	0.85	0.79	0.62	0.49
Ours	0.86	0.80	0.91	0.89	0.93	0.89	0.81	0.80

low recognition efficiency. This problem can be adjusted during data preprocessing. Referring to different methods is note recognition efficiency, our method note detail information recognition efficiency is kept above 80%, and our method is significantly better than other deep learning methods.

5. Conclusion

Music teaching is often difficult to grasp the characteristics of the notes and tones of the musical notation, and traditional teaching methods do not allow students to have a comprehensive understanding of the notation. This reduces the efficiency of music teaching. To improve the quality and efficiency of music teaching, we try to automate the teaching of music notation. With the addition of computer vision technology and note recognition algorithms, we improve the generative adversarial network to enhance the recognition accuracy and efficiency of music short scores. We adopt an embedded matching structure based on adversarial neural networks, starting from generators and discriminators, respectively, to unify generators and discriminators from the note input side. Each network layer is then laid out according to a cascade structure to preserve the different layers of note features in each convolutional layer. Residual blocks are then inserted in some network layers to break the symmetry of the network structure and enhance the ability of the adversarial network to acquire note features. To

validate the efficiency of our method, we selected the monophonic spectral dataset Bach Chorales, the polyphonic spectral dataset Video Game, and the miscellaneous spectral dataset Lakh for validation. The experimental results prove that our method has the best recognition accuracy in both monophonic and miscellaneous phonetic spectra, and in the recognition efficiency of note detail information, our method maintains more than 80%, which is better than other deep learning methods.

Compared with machine learning methods and deep learning methods, our method still has much room for improvement in recognition accuracy and recognition efficiency, although it performs best in the music notation recognition experiments. In future research, we will try to add recurrent neural networks as auxiliary classification in the adversarial network to optimize the recognition of mixed notes during note segmentation and improve the robustness and generalization of the network.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] L. B. Teen, A. b H. Ramli, and L. F. Ying, "Raising interest with software integration in music notation learning," *Procedia-Social and Behavioral Sciences*, vol. 143, pp. 379–383, 2014.
- [2] P. M. Bissell, "Tune in to technology," *Music Educators Journal*, vol. 85, no. 2, pp. 36–41, 1998.
- [3] R. Xiao, H. Tang, P. Gu, and X Xu, "Spike-based encoding and learning of spectrum features for robust sound recognition," *Neurocomputing*, vol. 313, pp. 65–73, 2018.
- [4] M. Simonović, M. Kovandžić, I. Ćirić, and V. Nikolic, "Acoustic recognition of noise-like environmental sounds by using artificial neural network," *Expert Systems with Applications*, vol. 184, Article ID 115484, 2021.
- [5] D. S. Naser and G. Saha, "Influence of music liking on EEG based emotion recognition," *Biomedical Signal Processing and Control*, vol. 64, Article ID 102251, 2021.
- [6] J. Calvo-Zaragoza, A. H. Toselli, and E. Vidal, "Handwritten Music Recognition for Mensural notation with convolutional recurrent neural networks," *Pattern Recognition Letters*, vol. 128, pp. 115–121, 2019.
- [7] J. Calvo-Zaragoza and J. Oncina, "Recognition of pen-based music notation with finite-state machines," *Expert Systems with Applications*, vol. 72, pp. 395–406, 2017.
- [8] M. S. Sinith, S. Tripathi, and K. V. V. Murthy, "Raga recognition using fibonacci series based pitch distribution in Indian Classical Music," *Applied Acoustics*, vol. 167, Article ID 107381, 2020.
- [9] S. I. Muzaffar, K. Shahzad, F. Aslam, M. Khalid, and K. Malik, "Process matching: performance trade-off between summary and full-length descriptions," *Computing and Informatics*, vol. 38, no. 4, pp. 851–882, 2019.
- [10] C. Wen, A. Rebelo, J. Zhang, and J. Cardoso, "A new optical music recognition system based on combined neural network," *Pattern Recognition Letters*, vol. 58, pp. 1–7, 2015.

- [11] C. Wolf, J. M. Jolion, and F. Chassaing, "Text localization, enhancement and binarization in multimedia documents," *Object recognition supported by user interaction for service robots*. IEEE, vol. 2, pp. 1037–1040, 2002.
- [12] B. Moysset, C. Kermorvant, C. Wolf, and J. Louradour, "Paragraph Text Segmentation into Lines with Recurrent Neural networks," in *Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 456–460, IEEE, Tunis, Tunisia, August 2015.
- [13] F. D. Julca-Aguilar and N. S. T. Hirata, "Image operator learning coupled with CNN classification and its application to staff line removal," in *Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, pp. 53–58, IEEE, Sao Paulo, Brazil, September 2017.
- [14] Q. N. Vo, S. H. Kim, H. J. Yang, and G Lee, "Binarization of degraded document images based on hierarchical deep supervised network," *Pattern Recognition*, vol. 74, pp. 568–586, 2018.
- [15] Y. Xu, R. Jia, L. Mou et al., "Improved Relation Classification by Deep Recurrent Neural Networks with Data augmentation," 2016, <https://arxiv.org/abs/1601.03651>.
- [16] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding Data Augmentation for Classification: When to warp?" in *Proceedings of the 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pp. 1–6, IEEE, Gold Coast, Australia, December 2016.
- [17] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: a retrospective," *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98–136, 2015.
- [18] J. Dai, Y. Li, K. He, and J. Sun, "Object detection via region-based fully convolutional networks," *Advances in Neural Information Processing Systems*, vol. 29, 2016.
- [19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards real-time object detection with region proposal networks[J]," *Advances in Neural Information Processing Systems*, vol. 28, 2015.
- [20] A. Pacha, J. Hajič, and J. Calvo-Zaragoza, "A baseline for general music object detection with deep learning," *Applied Sciences*, vol. 8, no. 9, p. 1488, 2018.
- [21] A. Baró, P. Riba, J. Calvo-Zaragoza, and A. Fornes, "From optical music recognition to handwritten music recognition: a baseline," *Pattern Recognition Letters*, vol. 123, pp. 1–8, 2019.
- [22] J. Hajic, M. Dorfer, G. Widmer, and P. Pecina, "Towards Full-Pipeline Handwritten OMR with Musical Symbol Detection by u-nets," 2018, <https://archives.ismir.net>.
- [23] L. Pugin, "Optical Music Recognitoin of Early Typographic Prints Using Hidden Markov Models," in *Proceedings of the 2006, 7th International Conference on Music Information Retrieval*, pp. 53–56, ISMIR, Victoria, Canada, October 2006.
- [24] J. Calvo-Zaragoza, A. H. Toselli, and E. Vidal, "Early Handwritten Music Recognition with Hidden Markov models," in *Proceedings of the 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 319–324, IEEE, Shenzhen, China, October 2016.
- [25] J. Calvo-Zaragoza, A. H. Toselli, and E. Vidal, "Handwritten music recognition for mensural notation: formulation, data and baseline results," in *Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1081–1086, IEEE, Kyoto, Japan, November 2017.
- [26] D. Amodei, S. Ananthanarayanan, R. Anubhai et al., "Deep speech 2: end-to-end speech recognition in English and Mandarin[C]//International conference on machine learning," in *Proceedings of the 33rd International Conference on International Conference on Machine Learning PMLR*, vol. 48, pp. 173–182, New York, USA, June 2016.
- [27] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2298–2304, 2017.
- [28] J. Calvo-Zaragoza and D. Rizo, "End-to-End neural optical music recognition of monophonic scores," *Applied Sciences*, vol. 8, no. 4, p. 606, 2018.
- [29] E. van Der Wel and K. Ullrich, "Optical Music Recognition with Convolutional Sequence-To-Sequence models," 2017, <https://arxiv.org/abs/1707.04877>.
- [30] Y. H. Lai and S. H. Lai, "Emotion-preserving Representation Learning via Generative Adversarial Network for Multi-View Facial Expression recognition," in *Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pp. 263–270, IEEE, Xi'an, China, May 2018.
- [31] H. Yang, U. Ciftci, and L. Yin, "Facial Expression Recognition by De-expression Residue learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2168–2177, Salt Lake City, UT, USA, June 2018.
- [32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [33] Y. Li, R. Bu, M. Sun, W. Wei, D. Xinhan, and C. Baoquan, "Pointcnn: convolution on x-transformed points," *Advances in Neural Information Processing Systems*, vol. 5, p. 31, 2018.
- [34] I. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," *Advances in Neural Information Processing Systems*, vol. 27, 2014.
- [35] R. Huang, S. Zhang, T. Li, and H. Ran, "Beyond face rotation: global and local perception gan for photorealistic and identity preserving frontal view synthesis," in *Proceedings of the IEEE international conference on computer vision*, pp. 2439–2448, Venice, Italy, October 2017.
- [36] G. Hadjeres, F. Pachet, and F. Nielsen, "Deepbach: a steerable model for bach chorales generation," in *Proceedings of the 34th International Conference on Machine Learning*, vol. 70, pp. 1362–1371, Sydney, Australia, August 2017.
- [37] F. T. Liang, M. Gotham, M. Johnson, and J. Shotton, "Automatic stylistic composition of bach chorales with deep LSTM," in *Proceedings of the In 18th International Society for Music Information Retrieval ConferenceISMIR*, pp. 449–456, Cambridge, UK, October 2017.
- [38] C. Z. A. Huang, C. Hawthorne, A. Roberts et al., "The bach doodle: approachable music composition with machine learning at scale," 2019, <https://arxiv.org/abs/1907.06637>.
- [39] C. Raffel and D. P. Ellis, "Extracting ground-truth information from MIDI files: a MIDIfesto," pp. 796–802, 2016, <https://www.ismir.net>.
- [40] C. Z. A. Huang, A. Vaswani, J. Uszkoreit et al., "Music transformer: generating music with long-term structure," 2018, <https://arxiv.org/abs/1809.04281>.

Research Article

Prediction of Diabetes through Retinal Images Using Deep Neural Network

Mahmoud Ragab ^{1,2,3}, **Abdullah S. AL-Malaise AL-Ghamdi** ^{4,5,6}, **Bahjat Fakieh** ⁴,
Hani Choudhry ^{2,7}, **Romany F. Mansour** ⁸, and **Deepika Koundal** ⁹

¹Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

²Centre for Artificial Intelligence in Precision Medicines, King Abdulaziz University, Jeddah 21589, Saudi Arabia

³Mathematics Department, Faculty of Science, Al-Azhar University, Naser City 11884, Cairo, Egypt

⁴Information Systems Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁵Information Systems Department, HECI School, Dar Alhekkma University, Jeddah, Saudi Arabia

⁶Center of Excellence in Smart Environment Research, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁷Biochemistry Department, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁸Department of Mathematics, Faculty of Science, New Valley University, El-Kharga 72511, Egypt

⁹School of Computer Science, University of Petroleum & Energy Studies, Dehradun, India

Correspondence should be addressed to Deepika Koundal; dkoundal@ddn.upes.ac.in

Received 25 February 2022; Revised 19 April 2022; Accepted 17 May 2022; Published 3 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Mahmoud Ragab 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.

Microvascular problems of diabetes, such as diabetic retinopathy and macular edema, can be seen in the eye's retina, and the retinal images are being used to screen for and diagnose the illness manually. Using deep learning to automate this time-consuming process might be quite beneficial. In this paper, a deep neural network, i.e., convolutional neural network, has been proposed for predicting diabetes through retinal images. Before applying the deep neural network, the dataset is preprocessed and normalised for classification. Deep neural network is constructed by using 7 layers, 5 kernels, and ReLU activation function, and MaxPooling is implemented to combine important features. Finally, the model is implemented to classify whether the retinal image belongs to a diabetic or nondiabetic class. The parameters used for evaluating the model are accuracy, precision, recall, and F1 score. The implemented model has achieved a training accuracy of more than 95%, which is much better than the other states of the art algorithms.

1. Introduction

Diabetes is a condition in which the body's ability to process sugar (glucose) is impaired [1]. Because of this, glucose levels in the blood shoot through the roof. Hyperglycemia is the medical term for this condition [2]. The body is unable to create enough insulin when this occurs. There is also the potential that the body cannot respond to the produced insulin. Diabetes cannot be cured, but it can be managed. Nerve damage, heart attacks, kidney failure, and stroke are possible outcomes for diabetes. Diabetes affects an estimated

8.8% of the world's population, according to statistics from 2017 [3]. By 2045, this figure is expected to rise to 9.9 percent.

Type 1 diabetes (T1D) and type 2 diabetes are the two types of diabetes (T2D) [4]. Most people diagnosed with type 1 diabetes are in their teens or early twenties. High blood glucose levels and increased thirst and urination are the most common symptoms. Oral drugs alone are ineffective in treating this kind of diabetes; hence, insulin therapy is essential. Obesity, hypertension, dyslipidemia, arteriosclerosis, and other disorders are all more common in older adults and the elderly regarding type 2 diabetes [5].

Diabetes is becoming more and more widespread as people's standard of living rises. Diabetes diagnosis and analysis should be studied because of the importance of speed and accuracy. Glucose tolerance, fasting blood glucose levels, and random blood glucose levels are all used to diagnose diabetes in the medical community [6]. The sooner we get a diagnosis, the easier it will be to treat it. Based on a person's daily physical examination data, machine learning can assist humans in making a preliminary diagnosis of diabetes mellitus. The most critical issues in machine learning are identifying useful features and the correct classifier [7].

The standard machine learning methods, such as the support vector machine (SVM), the decision tree (DT), the logistic regression, and others, have recently been applied to predict diabetes [8]. PCA and fuzzy neural inference were used to separate patients with diabetes from those who were not. As a result of the QPSO method and weighted least squares support vector machine (WLS-SVM) developed by Chi et al. [9], type 2 diabetes can be predicted. Diabetes can be predicted using a model developed by Çalişir and Doğan- tekin, known as LDA-MWSVM [10]. Linear discriminant analysis (LDA) was utilized to reduce dimensionality and extract features in this system [11]. High-dimensional datasets necessitated logistic regression to build prediction models for diverse onsets of type 2 diabetes. SVR (support vector regression) was utilized by Georga et al. to predict diabetes, a multivariate regression problem [12]. To further enhance the accuracy of the results, a growing number of studies are using ensemble approaches [13]. Combining 30 machine learning algorithms, Ozcift and Gulden developed an ensemble approach known as rotation forest [14].

Diabetes can be accurately predicted using AI-based technologies [15]. Deciduous categorization power is one of the advantages of using decision trees in the medical industry. In addition, a random forest produces a large number of decision trees. Recently, neural networks have emerged as a well-known machine learning technique because of their superior overall performance. In this article, deep neural networks will be used to predict the onset of diabetes. The proposed work will use the convolutional neural network to indicate diabetes [16]. In addition, the proposed work has used the dataset consisting of retinal images, and a deep neural network will be implemented on this retinal image dataset to predict the disease [17].

The organisation of the paper is as follows: Section 1 describes the introduction of the paper, whereas Section 2 discusses about the background study. In Section 3, proposed methodology is explained with results and discussion in Section 5. Finally, the conclusion in Section 5 followed by the reference section.

2. Related Work

When it comes to saving a person's life, early diabetes diagnosis is critical. In the last several years, a number of new diagnostic methods have been developed based on various models and methodologies. Neural networks, deep learning, and machine learning are just a few of the methods that can be used to improve facial recognition [18]. Other methods

include decision making, KNN, retinal pictures, and face images for diagnosis [19].

Joshi and Borse [20] developed a neural network called back propagation (BPNN). MathWorks (MATLAB) was used to create the user-interface. Researchers utilise the Pima Indian Diabetes Dataset to test their proposed methods. Parsing is conducted after the dataset has been loaded. ANNs were trained using back propagation neural networks after reading the values one by one. During the feature extraction phase, values were grouped together based on shared characteristics, and the groups were then arranged in a table. In the proposed method, the following step was to normalise the data. The data were encoded as a binary number between 0 and 1 [21]. Data redundancy is eliminated, and data relationships are ensured, as a result of normalisation. The final phase in the proposed method was training. The proposed system underwent up to nine iterations of training. The third iteration yielded the lowest level of error. At lower epoch values, the best results were achieved. Regression and validation plots were used to generate the results but accuracy is not.

The computational speed and efficiency of feed forward ANN (FFANN) make it popular in today's society. Diagnosis of diabetes can be improved by using the Small World FANN model, according to Erkaymaz and Ozer [22]. Researchers considered a four-layered FFANN in their investigation. The network included one output neuron and eight inputs. In FFANN, they made use of two hidden layers. FFANN used two alternative network topologies. Scientists rely on a bipolar-sigmoid function as their activation function in developing the new approach. The SW-FFANN training algorithm was based on a backpropagation learning algorithm. The PIDD dataset from the University of California, Irvine (UCI) repository was used in this study. The drawback is that optimum regular topology for SW-network development had been used in the rewiring procedure, for which DGlobal and DLocal parameters were determined.

A technique based on artificial neural networks was described in detail. Input, hidden, and output layers make up the three main components of an artificial neural network. Raw data are sent to the input layer. Inputs and weights assigned to them determine how hidden layers work. The data were entered into a JNN tool that calculates the attributes' values. Training, testing, and validation of data were then carried out—Binary numbers were the output of the suggested system. As a diabetic, I scores 0 points, while a healthy one scores 1. The proposed system had an average error rate of 0.010. The dataset underwent a total of 158,000 iterations. There were 767 training samples and 237 validation samples. The limitation of this technique is that its computationally complex.

Aliberti et al. [23] examined the prediction algorithms trained on glucose signals from a large and heterogeneous cohort of patients and then applied them to estimate future glucose levels on a brand-new patient. Based on nonlinear autoregressive (NAR) and long short-term memory (LSTM) neural networks, the authors have developed and compared two different types of solutions that have been successful in numerous time series prediction situations [24].

A deep neural network framework based on stacked autoencoders was presented by Kannadasan et al. [25] to classify the diabetes data. First, stacked autoencoders are used to extract features from the dataset, and the dataset is then categorised using a softmax layer. Finally, the network is fine tuned using the training dataset using supervised backpropagation. Pradhan et al. [26] used skin impedance and heart rate variability to identify diabetes. Classification was accomplished with the usage of artificial neural networks. Six females and five males with diabetes, an average age of 8 to 40 years, had been studied for skin impedance data. In addition, data from eight normal people, five females, and three males, with an average age of 3 to 24 years, were gathered for the study. When it came to determining signal strength at various frequencies, the Welch method was employed. Data on the electrocardiograms of 20 healthy volunteers, 14 men, and 6 women, with an average age of 22 years and 7 months, were gathered. Additionally, information was gathered on 20 diabetes patients, eight of whom were female and twelve of whom were male, with an average age of 40 years and eight months. In order to remove baseline drift from the resulting signal, median filtering was employed during the initial stages of signal preprocessing. Butter worth a lowpass filter was also used to reduce the high-frequency noise. The Savitzky–Golay filter was then used to smooth the ECG signal [27].

A deep neural network screening model was created by Ryu et al. [28] for patients with undetected diabetes mellitus (DM). Data from the Korean National Health and Nutrition Examination b (KNHANES) from 2013–2016 were used in our cross-sectional investigation. Only 11,456 people were included in the study after removing those diagnosed with DM, those under the age of 20, and those with incomplete data. KNHANES 2013–2015 was utilised as a training dataset and evaluated to generate a DLM for undiagnosed diabetes mellitus. The DLM was tested on a sample of 4,444 people who completed the 2016 KNHANES survey. Age, waist circumference, BMI, gender, smoking status, hypertension, and family history of diabetes (FH) were used to build the DLM. The area under the curve (AUC) of the model was 80.11, which is in line with previous screening models' performance [29].

3. Materials and Methods

The proposed methodology has been discussed in this section. Multilayer neural networks have been employed in the suggested research as a deep NN [30]. Convolutional neural networks are gaining popularity as data are structured as an image. Normalization is a key part of this procedure, as it is used for most of the data. Before beginning any work, it is highly advised to preprocess the images from dataset. As a result, preprocessed data will help in improved accuracy. This dataset has been fed into proposed deep convolutional neural network after preprocessing and normalization. Deep neural networks (DNNs) are then used to run and fit our data, resulting in the output. The following sections will provide a high-level overview of the completed work. Figure 1 displays the flowchart of the proposed model.

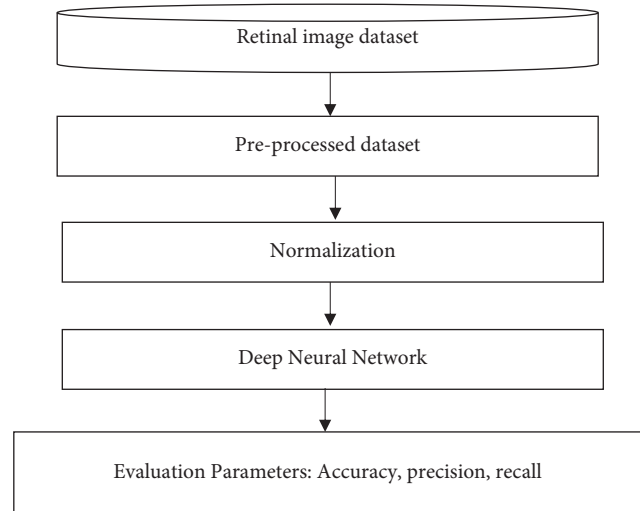


FIGURE 1: Flowchart of the proposed work.

3.1. Dataset. The images used for performing the analysis is downloaded from the Github <https://github.com/deepdrdoc/Deep-Diabetic-Retinopathy-Image-Dataset-DeepDRiD>. In this work, 410 images based on retina are used for performing the analysis and these images are classified for predicting the images as diabetic or nondiabetic.

3.2. Preprocessing Data. As previously stated, preprocessing is critical to this work. Image processing techniques are used to perform preprocessing on this dataset. To do this, the approaches described above have been employed to locate and bold the intensity of aberrant locations and parts. As a result, an unusual structure can be seen in some images. The optic disc and vessels, as an example, are not typical. The findings of previous tests using multichannel images were not very dependable; this was the case even before attempting a solution based on grayscale data. As a result, grayscale graphics were chosen as the preferred method of presentation. The next step is to normalize the photos after converting them to grayscale.

In the preprocessing stage, data may now be normalised easily by dividing image intensities to 255 (image converted to greyscale previously). The data must first be normalised using a label to form the network. Each picture name contains a class label in the first substring. The preprocessed data have been standardised as of this moment.

3.3. Creating Deep Neural Network Model. This section discusses deep neural networks with seven layers using different activation functions. For example, the first layer of the convolutional1D network uses the ReLU activation function with a kernel size of 5. After preprocessing, the data are reduced to a $256 * 256$ grayscale picture [31]. The input shape of size $(256, 256, 1)$ will be utilized as an input. After that, MaxPooling combines the most important features, then flattens the image, and finally, the classification will be done. The dense layer should be utilized because we have

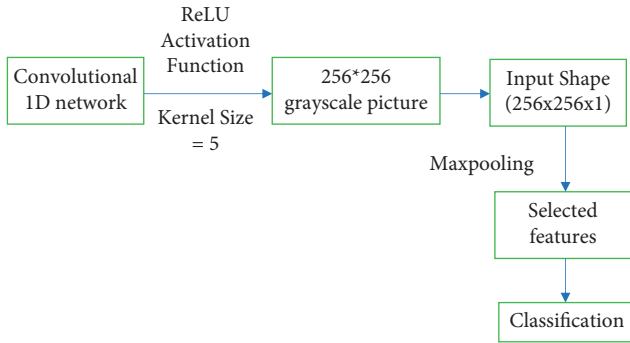
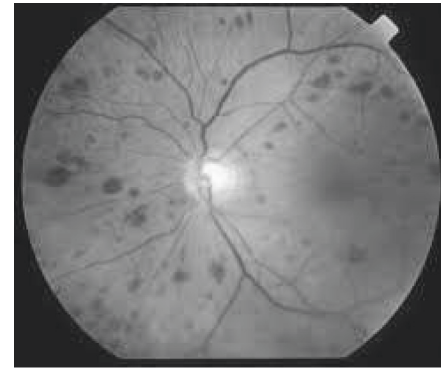


FIGURE 2: Architecture of deep neural network.



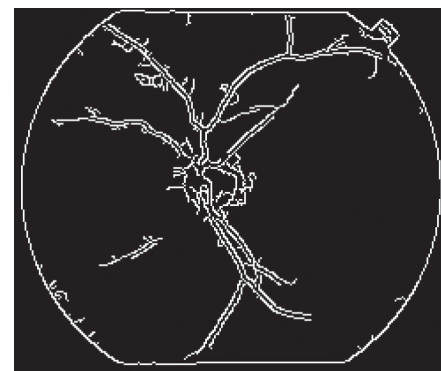
Normal Image

FIGURE 3: Normal image.



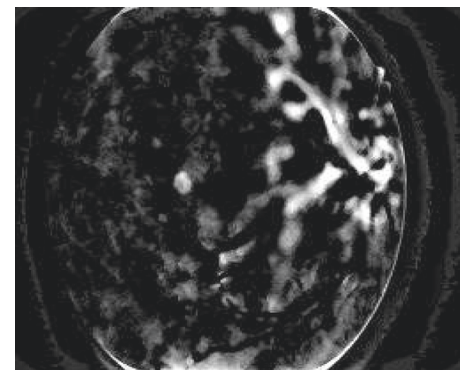
Gray Scaled Image

FIGURE 4: Gray scaled image.



Vessels detected by canny edge detection filter

FIGURE 5: Vessels detected by canny edge detection filter.



Preprocessed image sent to the CNN. white_top_hat

FIGURE 6: Preprocessed image sent to the CNN. white_top_hat + gray_scaled.

binary categories (diabetic vs. nondiabetic). Because our class labels are binary, our loss function is binary cross-entropy. Finally, Adam’s batch size is set to 10, and it is used as an optimizer in batch mode. This optimizer will help to prevent the overfitting of the proposed model [32].

Figure 2 has shown the architecture of the proposed neural network with the size representation of the images. The Figure 3 has also shown the use of activation function and the max pooling layer for performing the classification.

4. Results and Discussion

The proposed work is implemented using the anaconda framework for executing the Python codes. Anaconda framework is rich in libraries based on deep learning due to which TensorFlow and Keras were easily imported in the code. Different aspects of the image taken for the analysis are given below:

Figures 3–6 has shown the different shades of the images. Figure 3 has represented the normal image, whereas Figure 4 has illustrated the gray scaled image. Figure 5 has shown the vessels detected by canny edge detection filter, and at last, Figure 6 has shown the generated preprocessed image which is further classified by the deep neural network.

Above Tables 1 and 2 has shown the comparison chart and confusion matrix of the proposed work, respectively. In

Table 1, the proposed deep neural network is compared with the existing machine learning algorithms and the results have indicated that the proposed work has performed better than the existing machine learning algorithms.

TABLE 1: Comparison chart of the proposed work.

Model	Class-label	Precision	Recall	F1-score	Accuracy (%)
Logistic regression [33]	0 (nondiabetic)	0.72	0.71	0.93	73
	1 (diabetic)	0.73	0.72	0.94	
Random forest	0 (nondiabetic)	0.76	0.75	0.75	77.4
	1 (diabetic)	0.75	0.77	0.76	
Proposed fine-tuned MLP	0 (nondiabetic)	0.86	0.85	0.85	86.6
	1 (diabetic)	0.86	0.88	0.87	
Deep neural network (proposed)	0 (nondiabetic)	0.95	0.91	0.93	95.6
	1 (diabetic)	0.94	0.93	0.94	

TABLE 2: Confusion matrix.

	True	Positive
False	0.95	0.91
True	0.94	0.93

5. Conclusion

Diagnosing diabetes at an early stage is critical to finding an effective treatment. Diabetes classification is implemented using a deep neural network, i.e., convolutional neural network, in the current work. The dataset contains more than 410 images based on the retina for diabetes classification. The number of training epochs was kept short of ensuring that the approach could be quickly used on any mobile device. The experimental results suggest that the proposed deep learning model is effective and accurate. The model has achieved an accuracy greater than 95%. The model for determining all probable complications, including an orderly sequence in terms of the percentage of complications that can occur, will be improved in a future study as well. Additionally, deep learning algorithms and methodologies can be incorporated to enhance the work for an automated diabetes analysis.

Data Availability

Data are publicly available at <https://github.com/deepdrdoc/Deep-Diabetic-Retinopathy-Image-Dataset-DeepDRiD>.

Conflicts of Interest

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

Acknowledgments

This work was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant no. (D-888-611-1443). The authors, therefore, gratefully acknowledge DSR for technical and financial support.

References

- [1] X. Zhao, Y. Liu, W. Zhang et al., "Relationships between retinal vascular characteristics and renal function in patients with type 2 diabetes mellitus," *Translational Vision Science & Technology*, vol. 10, no. 2, p. 20, 2021.
- [2] M. Björk, E. O. Melin, T. Frisk, and M. Thunander, "Admission glucose level was associated with increased short-term mortality and length-of-stay irrespective of diagnosis, treating medical specialty or concomitant laboratory values," *European Journal of Internal Medicine*, vol. 75, pp. 71–78, 2020.
- [3] P. Saeedi, I. Petersohn, P. Salpea et al., "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the international diabetes federation diabetes atlas, 9th edition," *Diabetes Research and Clinical Practice*, vol. 157, Article ID 107843, 2019.
- [4] S. Tsalamandris, A. S. Antonopoulos, E. Oikonomou et al., "The role of inflammation in diabetes: current concepts and future perspectives," *European Cardiology Review*, vol. 14, no. 1, pp. 50–59, 2019.
- [5] J. Freeman, "Management of hypoglycemia in older adults with type 2 diabetes," *Postgraduate Medicine*, vol. 131, no. 4, pp. 241–250, 2019.
- [6] M. Hod, "Gestational diabetes and new diagnostic criteria," *Journal of Perinatal Medicine*, vol. 41, 2013.
- [7] R. Ferdousi, M. A. Hossain, and A. el Saddik, "Early-stage risk prediction of non-communicable disease using machine learning in Health CPS," *IEEE Access*, vol. 9, pp. 96823–96837, 2021.
- [8] N. S. Prema, V. Varshith, and J. Yogeswar, "Prediction of diabetes using ensemble techniques," *International Journal of Recent Technology and Engineering*, vol. 7, no. 6, 2019.
- [9] C. Yue, L. Xin, X. Kewen, and S. Chang, "An intelligent diagnosis to type 2 diabetes based on QPSO algorithm and WLS-SVM," in *Proceedings of the 2008 International Symposium on Intelligent Information Technology Application Workshops*, Shanghai, China, December 2008.
- [10] D. Çalışır and E. Doğantekin, "An automatic diabetes diagnosis system based on LDA-Wavelet Support Vector Machine Classifier," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8311–8315, 2011.
- [11] R. Mehta, K. Aggarwal, D. Koundal, A. Alhudhaif, and K. Polat, "Markov features based DTCWS algorithm for online image forgery detection using ensemble classifier in the pandemic," *Expert Systems with Applications*, vol. 185, Article ID 115630, 2021.
- [12] E. I. Georga, V. C. Protopappas, D. Ardigò, D. Polyzos, and D. I. Fotiadis, "A glucose model based on support vector regression for the prediction of hypoglycemic events under free-living conditions," *Diabetes Technology & Therapeutics*, vol. 15, no. 8, pp. 634–643, 2013.
- [13] R. I. Doewes, R. Nair, and T. Sharma, "Diagnosis of COVID-19 through blood sample using ensemble genetic algorithms and machine learning classifier," *World Journal of Engineering*, vol. 19, no. 2, pp. 175–182, 2021.

- [14] A. Ozcift and A. Gulten, "Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms," *Computer Methods and Programs in Biomedicine*, vol. 104, no. 3, pp. 443–451, 2011.
- [15] S. Aggarwal, S. Gupta, A. Alhudhaif, D. Koundal, R. Gupta, and K. Polat, "Automated COVID-19 detection in chest X-ray images using fine-tuned deep learning architectures," *Expert Systems*, vol. 39, no. 3, 2021.
- [16] R. Nair, R. Alhudhaif, D. Koundal, R. I. Doewes, and P. Sharma, "Deep learning-based COVID-19 detection system using pulmonary CT scans," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 29, no. SI-1, pp. 2716–2727, 2021.
- [17] R. Krishnamoorthi, S. Joshi, H. Z. Almarzouki et al., "A novel diabetes healthcare disease prediction framework using machine learning techniques," *Journal of Healthcare Engineering*, vol. 2022, Article ID 1684017, 10 pages, 2022.
- [18] D.-J. Chiu, "Deep learning based automated detection of diseases from apple leaf images," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1849–1866, 2022.
- [19] R. Nair, S. Vishwakarma, M. Soni, T. Patel, and S. Joshi, "Detection of COVID-19 cases through X-ray images using hybrid deep neural network," *World Journal of Engineering*, vol. 19, no. 1, pp. 33–39, 2021.
- [20] S. Joshi and M. Borse, "Detection and prediction of diabetes mellitus using back-propagation neural network," in *Proceedings of the 2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)*, Ghaziabad, Uttarpradesh, September 2016.
- [21] S. Shambhu, D. Koundal, P. Das, and C. Sharma, "Binary classification of COVID-19 CT images using CNN," *International Journal of E-Health and Medical Communications*, vol. 13, no. 2, pp. 1–13, 2022.
- [22] O. ErKaymaz and M. Ozer, "Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes," *Chaos, Solitons & Fractals*, vol. 83, pp. 178–185, 2016.
- [23] A. Aliberti, I. Pupillo, S. Terna et al., "A multi-patient data-driven approach to blood glucose prediction," *IEEE Access*, vol. 7, pp. 69311–69325, 2019.
- [24] H. Z. Almarzouki, H. Alsulami, A. Rizwan, M. S. Basingab, H. Bukhari, and M. Shabaz, "An internet of medical things-based model for real-time monitoring and averting stroke sensors," *Journal of Healthcare Engineering*, vol. 2021, Article ID 1233166, 9 pages, 2021.
- [25] K. Kannadasan, D. R. Edla, and V. Kuppili, "Type 2 diabetes data classification using stacked autoencoders in deep neural networks," *Clinical Epidemiology and Global Health*, vol. 7, no. 4, pp. 530–535, 2019.
- [26] N. Pradhan, G. Rani, V. S. Dhaka, and R. C. Poonia, "Diabetes prediction using artificial neural network," in *Deep Learning Techniques for Biomedical and Health Informatics*, pp. 327–339, Academic Press, Massachusetts, MA, USA, 2020.
- [27] A. Hasan, A. Khan, A. Parveen, and R. Nair, "Emotion prediction through EEG recordings using computational intelligence," in *Computational Intelligence for Information Retrieval*, pp. 47–61, CRC Press, Boca Raton, FL, USA, 2021.
- [28] K. S. Ryu, S. W. Lee, E. Batbaatar, J. W. Lee, K. S. Choi, and H. S. Cha, "A deep learning model for estimation of patients with undiagnosed diabetes," *Applied Sciences*, vol. 10, no. 1, p. 421, 2020.
- [29] B. Sharma, D. Koundal, and S. Singh, "Medical imaging security and forensics: a systematic literature review," in *Advances in Computational Techniques for Biomedical Image Analysis*, pp. 273–297, Elsevier, Amsterdam, Netherlands, 2020.
- [30] R. Nair and A. Bhagat, "Genes expression classification through histone modification using temporal neural network," *Recent Advances in Computer Science and Communications*, vol. 14, no. 5, pp. 1488–1496, 2021.
- [31] S. A. Alex, J. J. V. Nayahi, H. Shine, and V. Gopirekha, "Deep convolutional neural network for diabetes mellitus prediction," *Neural Computing & Applications*, vol. 34, no. 2, pp. 1319–1327, 2022.
- [32] S. K. Jangir, N. Joshi, M. Kumar, D. K. Choubey, S. Singh, and M. Verma, "Functional link convolutional neural network for the classification of diabetes mellitus," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 37, no. 8, 2021.
- [33] U. M. Butt, S. Letchmunan, M. Ali, F. H. Hassan, A. Baqir, and H. H. R. Sherazi, "Machine learning based diabetes classification and prediction for healthcare applications," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9930985, 17 pages, 2021.

Research Article

The Collection and Recognition Method of Music and Dance Movement Based on Intelligent Sensor

Jun Geng 

School of Art, Shandong University of Finance and Economics, Jinan 250014, China

Correspondence should be addressed to Jun Geng; 20040293@sdufe.edu.cn

Received 22 February 2022; Revised 18 March 2022; Accepted 24 March 2022; Published 3 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Jun Geng. 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 popularization and development of Internet of Things technology, a large number of music and dance videos have emerged in all walks of life. In this information age, video communication has become a widespread communication method. In the current music and dance collection process, most of the action frame information of the dance video is repeated, and the stage background and costumes of the dance action are too many to fully express the human body movement information. Based on these problems, this article will realize the application of the intelligent sensor-based action recognition technology in the field of dance movement collection and complete the collection and recognition of music and dance movements. The research results of the article show that: (1) in the dance video image extraction process, the feature recognition effect of the proposed algorithm is the highest among the three models. The recognition effect of the upper body is 66.1, and the recognition effect of the lower body is 61.0. The image recognition effect can reach 73.4. During the statistical experiments on the recognition of different regions of the human body, the recognition effect of the intelligent sensor model proposed in the article is still the highest among the three models. The recognition effect of the upper body is 33.9, and the recognition effect of the lower body is 33.9. The recognition effect is 34.5, and the recognition effect of the whole body is 40.7. (2) In the traditional music and dance collection mode, the P values of the four test parts are all greater than 0.05, indicating that in the traditional music and dance collection mode, the differences between the four test modules are not significant. Combined with the evaluation results of the three groups in the traditional music and dance collection mode, we can conclude that under the condition that the initial conditions are basically the same, and the training conditions and environment are basically the same, the trainees who use the smart sensor music and dance collection training method are better in physical fitness. The indicators have been better improved, and the effect is greatly optimized compared with the training effect in the traditional music and dance collection mode. (3) After the test set runs, the article proposes that the accuracy rate of the dance collection model based on the smart sensor algorithm is 88.24%, the accuracy rate can reach 88.96%, the improved accuracy rate can reach 91.46%, and the accuracy rate can reach 91.79%. The ROC curve value of the article and the improved model is very stable. The ROC value before the improvement remains at about 0.90, and the ROC value after the model improvement also remains at 0.96. After the test set runs, the performance of the four models has decreased to a certain extent, but the smart sensor dance acquisition model proposed in the article has the lowest degree of decline, and the performance after the decline is still the highest among the four models. The accuracy of the model is 90.24%, and the accuracy of the improved model is 93.16%. The ROC curve values of the improved system are very stable, the ROC value has been maintained at 0.95, and the ROC value before the improvement is stable within the range of 0.85–0.95. The experimental results further illustrate that the model proposed in the article has the best performance.

1. Introduction

Dance is a kind of traditional culture with the beauty and significance of movement. We must carry forward it. Dance is composed of many movements. One of the challenges of dance development is the effective connection between the movements. In recent years, music and dance video analysis has also

become the research direction of many experts. Effective analysis of music and dance videos can not only correct dancers' posture, help dancers effectively train, but also play a certain role in the development of music and dance. *Promotion*. In view of the complexity in the dance learning process, the article decomposes complex dance videos into many small segments based on intelligent sensor technology, making the

abstract content more intuitive and easy to understand. Reference [1] explored the way of perceiving the correspondence between music and body movement by analyzing the characteristic relationship between music and movement. Literature [2] studies collections with modern dance with the aim of providing libraries with an annotated bibliographic resource guide. Reference [3] proposed a gesture recognition algorithm for Indian classical dance style using Kinect sensor, which can achieve a high recognition rate of 86.8%. Reference [4] introduced the meaning of Laban movement analysis symbols and the method of dance movement recognition. Reference [5] proposes a gesture recognition implementation that uses hidden Markov models to classify specific gestures in traditional dance. Reference [6] proposed a fuzzy membership function to automatically recognize dance gestures. Reference [7] introduces the gesture recognition method and its application in the dance training system in the teaching virtual reality environment. Reference [8] applied the recognition of dance gestures to music performances, reversing the traditional music and dance interaction. Reference [9] builds a robust recognizer based on linguistic motivation method to recognize dance poses in Balinese traditional dance choreography. Reference [10] presents a complete manipulation prototype for compressing and recognizing dance gestures in contemporary ballet. Reference [11] presents an update of a project aimed at developing a modular system for real-time analysis of body movements and postures. Reference [12] proposes a tool for nonuniform subsampling of spatiotemporal signals with the goal of recognizing a set of dance gestures in contemporary ballet. In [13], an angle representation of the skeleton is designed for robustness under the noise input of the Kinect sensor. Reference [14] designed a 4-stage system for automatic gesture recognition in dance. Reference [15] proposed a new method for human gesture recognition based on quadratic curve, which achieved a recognition rate of up to 97.65% on a database of 16 different gestures.

2. Research on the Collection and Recognition Methods of Music and Dance Movements

2.1. Motion Capture Classification. After consulting relevant materials, we found out the common dance video capture systems. The details are shown in Table 1.

2.2. Design of Dance Gesture Capture. The research of this paper is based on the principle of intelligent sensor and applies motion capture technology to the process of dance pose extraction, so as to extract the skeleton movement route of the human body and establish a quantitative method for dance pose analysis. First, install data identification points on each part of the dance trainer's body, and then dance in a predetermined area, and the system will automatically capture the basic dance movements of the dancer. In this process, the system will match the collected dance movements with

the dance movement database. When all the data collection points are successfully identified by the computer and entered into the system, the real-time collection of human dance movements has been successfully completed. The working roadmap of the scheme is shown in Figure 1.

3. Collection and Recognition of Music and Dance Movements by Smart Sensors

3.1. Dance Image Preprocessing. In this paper, the average value method is selected to grayscale the original image. This method obtains the gray value of the point by calculating the average value of the 3-channel brightness of each pixel point [18]. The corresponding mathematical expression is as follows:

$$\text{Gray} = \frac{(R + G + B)}{3}. \quad (1)$$

Assuming that the value of a certain pixel of the video at time is M_t , the random probability [19] is as follows:

$$P(M_t) = \sum_{i=1}^k \omega_{i,t} g(M_t, \mu_{i,t}, \sigma_{i,t}), \quad (2)$$

where K is the number of Gaussian distributions, ω_i is the i -th Gaussian distribution weight, $\mu_{i,t}$ and $\sigma_{i,t}$ represent the mean and variance, and g is the Gaussian distribution function.

$$|M_{t+1} - \mu_{i,t}| \leq 2.5\delta_{i,t}. \quad (3)$$

Parameter update is as follows:

$$\begin{aligned} \omega_{i,t+1} &= (1 - \rho)\omega_{i,t} + \rho, \\ \mu_{i,t+1} &= (1 - \alpha)\mu_{i,t} + \alpha M_t, \\ \delta_{i,t+1} &= (1 - \alpha)\delta_{i,t} + \alpha (M_t - \mu_{i,t})^T (M_t - \mu_{i,t}), \\ \alpha &= \rho g(M_t, \mu_{i,t}, \sigma_{i,t}), \end{aligned} \quad (4)$$

where ρ is the learning rate of the model, which ranges from 0 to 1, and α is the parameter update factor, which determines the speed of the update.

After the update is complete, use the value of $\omega_{i,t}/\delta_{i,t}$ to sort the K Gaussian distributions

$$B = \operatorname{argmin} \left(\sum_{k=1}^b \omega_k > T \right). \quad (5)$$

The effect of background subtraction on the gray-scaled dance video image using the Gaussian mixture model is shown in Figure 2:

3.2. Feature Extraction of Dance Images. Optical flow calculation can accurately calculate the changes of dance movements without missing the movement changes of small limbs. The calculation formula is as follows:

TABLE 1: Classification statistics of motion capture methods.

Capture method	Structure	Advantage	Disadvantage
Mechanical motion capture	Rigid rod attached to the body	Low cost and less restrictive	Movement is restricted, operations are complex, and data processing is complicated
Optical motion capture	High-precision cameras plus data collection points for joints [16]	Low movement restrictions, low cost, and high accuracy	The later data processing is huge, and the signal requirements are high
Electromagnetic motion capture	Electromagnetic generating equipment and receiving equipment	Inexpensive and fast response	More sensitive to the environment
Video motion capture	Camera and image processing device	High cost and low environmental requirements	The image algorithm is responsible, and the implementation is difficult
Acoustic motion capture	Acoustic emission source and receiving processing device [17]	Low price and high environmental adaptability	Poor accuracy and susceptible to interference

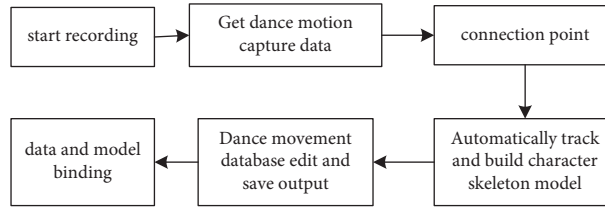


FIGURE 1: The process of acquiring dance gesture data.

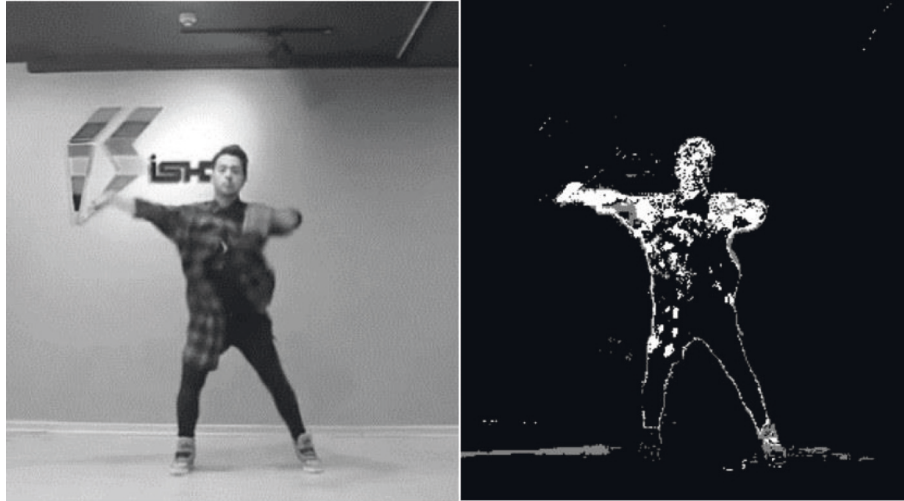


FIGURE 2: Background removal result.

$$\begin{aligned}
 E(w) &= E_{\text{color}}(w) + \gamma E_{\text{grad}}(w) + \alpha E_{\text{smooth}}(w) \\
 &\quad + \beta E_{\text{match}}(w, w_1) + E_{\text{desc}}(w_1), \\
 E_{\text{color}}(w) &= \int_{\Omega} \Psi(|I_2(x+w(x)) - I_1(x)|^2) dx, \\
 E_{\text{grad}}(w) &= \int_{\Omega} \Psi(|\nabla I_2(x+w(x)) - \nabla I_1(x)|^2) dx, \\
 E_{\text{smooth}}(w) &= \int_{\Omega} \Psi(|\nabla \mu(x)|^2 + |\nabla v(x)|^2) dx, \\
 E_E(w_1) &= \int \delta(x) |f_2(x+w_1(x)) - f_1(x)|^2 dx.
 \end{aligned} \tag{6}$$

The audio signal is stored in y , the length of y is N , the sampling rate is fs , and the formula for framing is as follows [20]:

$$fs = \frac{(N - \text{olap})}{\text{dis}} = \frac{(N - \text{wlen})}{\text{dis} + 1}. \tag{7}$$

Calculate the energy features of audio [21]

$$\begin{aligned}
 y_k(j) &= \text{win}(j) \cdot x((k-1) \cdot \text{dis} + j), 1 \leq j \leq L, 1 \leq k \leq f, \\
 M(k) &= \sum_{j=0}^{L-1} |y_k(j)|, 1 \leq k \leq f.
 \end{aligned} \tag{8}$$

Among the changing dance moves in a dance video, audio features can extract a representative frame

$$V = \frac{|H_{\text{current}} - H_{\text{key}}|}{H_{\text{key}}} \quad (9)$$

The similarity matching of human body poses is to realize the measure of the difference or similarity of the poses of different human bodies [19]

$$\begin{aligned} D &= \text{sqrt}((x_1 - x_2)^2 + (y_1 - y_2)^2), \\ D &= \text{sqrt}((x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2). \end{aligned} \quad (10)$$

Stop motion the image in the dance video, and then use the image block marker matrix to calculate the dance motion target, and get

$$L_{IB}(m, n) = \begin{cases} 2, & \text{SAD}(m, n) > t_{\text{SAD}}, \\ 1, & \text{IBSCI}(m, n) > t_{\text{IBSCI}}, \\ 0, & \text{else.} \end{cases} \quad (11)$$

Get the dance action gait contour target [22]

$$Z_k(x, y) = \begin{cases} 255, & |F_k(x, y) - B_k(x, y)| \geq \text{Th}, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

4. Test Experiments

4.1. Simulation Experiment. Human posture can be directly used as feature information for statistical action recognition. The experimental dance collection device is based on the intelligent sensor dance action database, which contains twenty groups of dance action combinations. The music and dance videos include two types of videos: one is a video of a combination of dance movements and the other is to collect data on the combined movements according to the type of movements. The video clip of each dance is about 20 seconds. The test set used in the experiment comes from the dance videos on the Internet. The intelligent sensor music and dance collection mode proposed in the article is compared with the recognition effect of the mechanical motion capture model and the optical motion capture model, and the recognition effect of the feature description is compared with different regions of the human body. The specific experimental data are as follows:

According to the experimental results in Table 2 and Figure 3, we can conclude that in the dance video image extraction process, the feature recognition effect of the algorithm proposed in this article is the highest among the three models. The recognition effect of the upper body is 66.1, the recognition effect of the lower body is 61.0, and the recognition effect of the whole image can reach 73.4. The recognition effect of optical motion capture is the lowest, the recognition effect of the whole image can only reach 60.4, and the recognition effect of mechanical motion capture is in the middle, and the recognition effect of the whole image is 69.1.

According to the data in Table 3 and Figure 4, we can conclude that in the process of statistical experiments on the recognition of different regions of the human body, the recognition effect of the intelligent sensor model proposed in this article is still the highest among the three models, and the recognition effect of the upper body is 33.9, the recognition effect of the lower body is 34.5, and the recognition effect of the whole body is 40.7. The full-body recognition effect of mechanical motion capture is 30.4, and the full-body recognition effect of optical motion capture is 22.4. According to the statistics of the two experimental results, it can also be shown that the recognition effect of the intelligent sensor dance acquisition model proposed in the article is optimal.

4.2. Comparative Experiment. In order to verify the effectiveness of the acquisition of music and dance movements by smart sensors on the effect of dance teaching, the experiment chose the method of comparative experiment, and the experiment selected the students who took aerobics as the research object in the 2019 grade of a university. In the experiment, 150 students who took aerobics as an elective were randomly divided into three groups: the routine group, the experimental group, and the training group. Before the experiment, a simple questionnaire survey was conducted on the members of the three groups to understand their learning motivation and learning efficiency. The experimental results showed that the P value was greater than 0.05, indicating that there was no significant difference between the three groups of students' learning motivation and learning efficiency. The experiment compares the students' dance performance of traditional music and dance collection methods and smart sensor music and dance collection methods, and observes the superiority of smart sensors in music and dance collection. In order to ensure the objectivity of the experimental data, the three groups of data were tested separately, and the main test contents were movement range, movement strength, movement consistency, movement specification, and 4 parts. The specific experimental data are as follows:

According to the data in Table 4 and Figure 5, we can conclude that under the traditional music and dance collection mode, the dance performance of the training group is the highest, with a range of movements of 80 points, strength of movements 77 points, consistency of movements 75 points, regularity of movements 74 points, and regularity points of 74 points. The dance score of the group was the lowest, with 76 points of movement range, 74 points of movement strength, 72 points of movement coherence, and 70 points of movement standardization. The experimental group's performance was in the middle of the two. In the traditional music and dance collection mode, the P values of the four test parts are all greater than 0.05, indicating that under the traditional music and dance collection mode, the differences between the four test modules are not significant.

According to the data in Table 5 and Figure 6, from the analysis of the evaluation results between the conventional group, the experimental group, and the training group, the overall situation of the students in the conventional training

TABLE 2: Comparison of recognition effects of feature description.

Body parts	Smart sensor	Mechanical motion capture	Optical motion capture
Upper body	66.1	60.9	52.8
Lower body	61.0	61.6	52.2
Whole body	67.1	55.7	53.3
Whole image	73.4	69.1	60.4

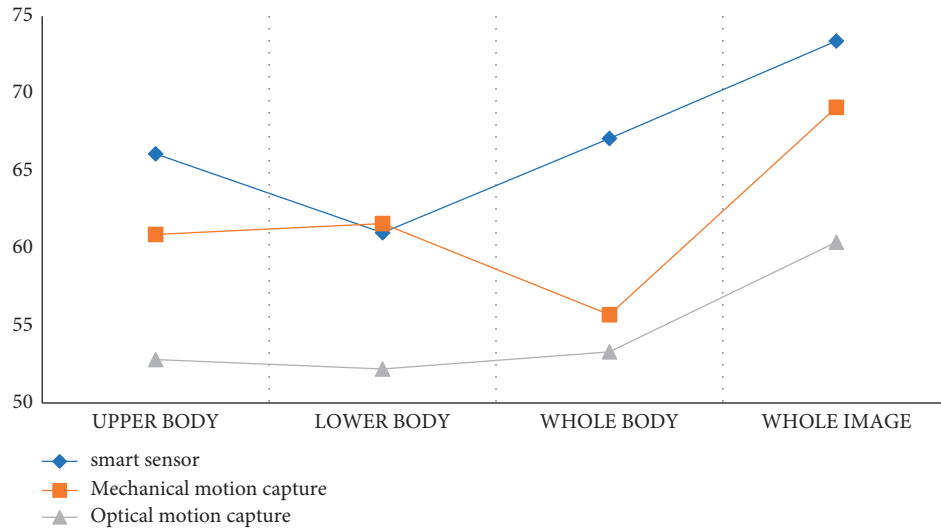


FIGURE 3: Feature description recognition effect statistics.

TABLE 3: Feature recognition effects of different regions of the human body.

Body parts	Smart sensor	Mechanical motion capture	Optical motion capture
Upper body	33.9	25.3	10.0
Lower body	34.5	25.8	10.3
Upper body + lower body	45.6	33.3	23.3
Whole body	40.7	30.4	22.4

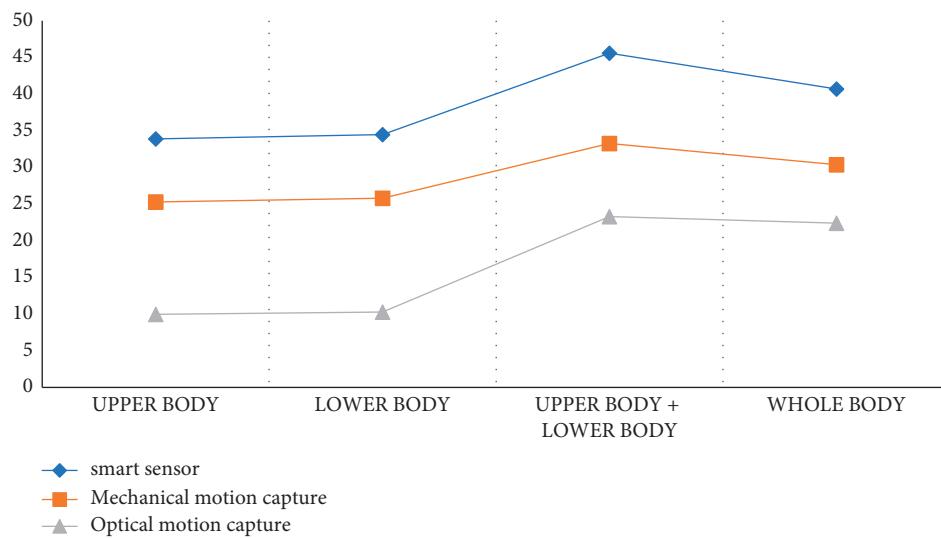


FIGURE 4: Feature recognition statistics of different regions of the human body.

TABLE 4: Collection methods of traditional music and dance.

Group	Range of motion	Action strength	Continuity of action	Action norm
Regular group	76	74	72	70
Test group	78	76	73	72
Training group	80	77	75	74
Significant P	0.42	0.33	0.39	0.43

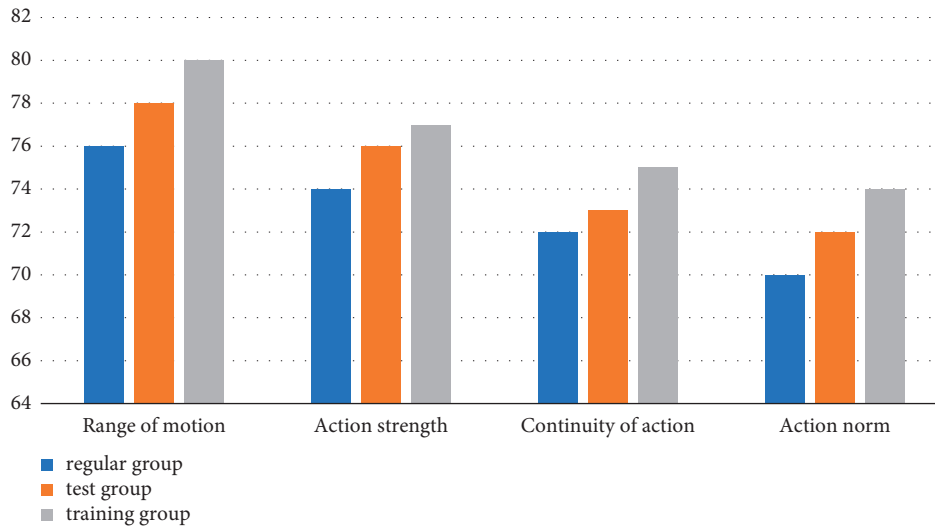


FIGURE 5: Collection statistics of traditional music and dance.

TABLE 5: Smart sensor music and dance collection methods.

Group	Range of motion	Action strength	Continuity of action	Action norm
Regular group	86	84	82	79
Test group	88	86	84	83
Training group	92	88	85	85
Significant P	0.42	0.28	0.23	<0.01

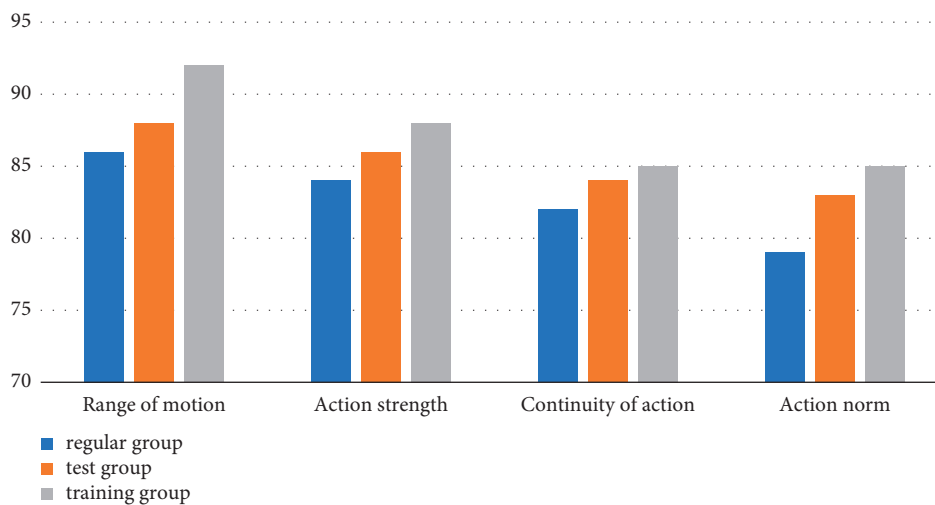


FIGURE 6: Smart sensor music and dance collection statistics.

group is slightly improved compared with the traditional music and dance collection mode. The P values of action coherence are all less than 0.05, indicating that there is a

large gap between the three. Among them, the P value of action normativeness is 0.05, which has a very significant difference. Compared with the traditional mode, the four

TABLE 6: Evaluation criteria table.

	Metrics	Formula
Accuracy	The accuracy measure refers to the ratio of the number of passing passes to all the numbers [23]. The larger the index value, the more accurate the detection result.	$\text{Precision} = \text{hits}_u / \text{recset}_u$
Recall	The recall criterion refers to the ratio of detections to the theoretical maximum hits [24]. The larger the index value, the more accurate the test result.	$\text{Recall} = \text{hits}_u / \text{testset}_u$
F1 measure	The F1 metric can effectively balance the precision and recall by biasing towards the side with a smaller value [25]. The larger the index value, the more accurate the test result.	$\text{F1} = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$

TABLE 7: The performance of each model on the test set.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Smart sensor dance collection model	88.24	88.96	89.30	89.48
Improved dance collection model	91.46	91.79	91.89	91.45
Mechanical motion capture model	84.13	84.43	84.79	85.19
Optical motion capture model	72.21	75.24	75.46	75.12

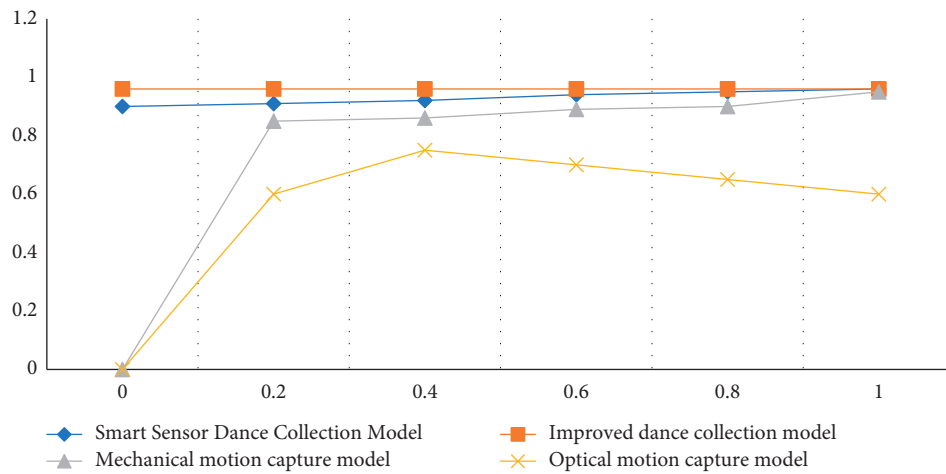


FIGURE 7: ROC curve under the test set.

test modules of the training group have improved effects. Combined with the evaluation results of the three groups in the traditional music and dance collection mode, we can conclude that under the condition that the initial conditions are basically the same, and the training conditions and environment are basically the same, the trainees who use the smart sensor music and dance collection training method are better in physical fitness. The indicators have been better improved, and the effect is greatly optimized compared with the training effect in the traditional music and dance collection mode.

4.3. Model Checking

4.3.1. Evaluation Criteria. The evaluation criteria are shown in Table 6.

4.3.2. Specific Tests. In order to test the performance of the proposed smart sensor-based music and dance acquisition model in music and dance acquisition, the experiment improved the proposed model and ran it on

the test set and the hybrid test set with mechanical motion capture and optical motion capture respectively. The test set is used to evaluate the generalization ability of the final model, and the mixed test set tunes the model's hyperparameters and is used to make an initial evaluation of the model's ability. ROC is a model evaluation indicator in the field of machine learning. The larger the ROC value of the classifier, the higher the accuracy rate. The specific experimental data are as follows:

According to the data in Table 7 and Figure 7, we can conclude that the accuracy rate of the dance acquisition model based on the smart sensor algorithm proposed in the article is 88.24%, the accuracy rate can reach 88.96%, the improved accuracy rate can reach 91.46%, and the accuracy rate can reach 91.79%, which is the highest index value among the four models in the experiment. The accuracy rate of the optical motion capture model is 72.21%, which is the lowest among the four models, and the mechanical motion capture model is in the middle state. According to the ROC curves of the four algorithms,

TABLE 8: The performance of each model on the mixed test set.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Smart sensor dance collection model	90.24	90.56	90.89	91.24
Improved dance collection model	93.16	93.89	93.78	94.40
Mechanical motion capture model	87.23	87.93	88.12	88.18
Optical motion capture model	75.14	75.24	75.89	7.12

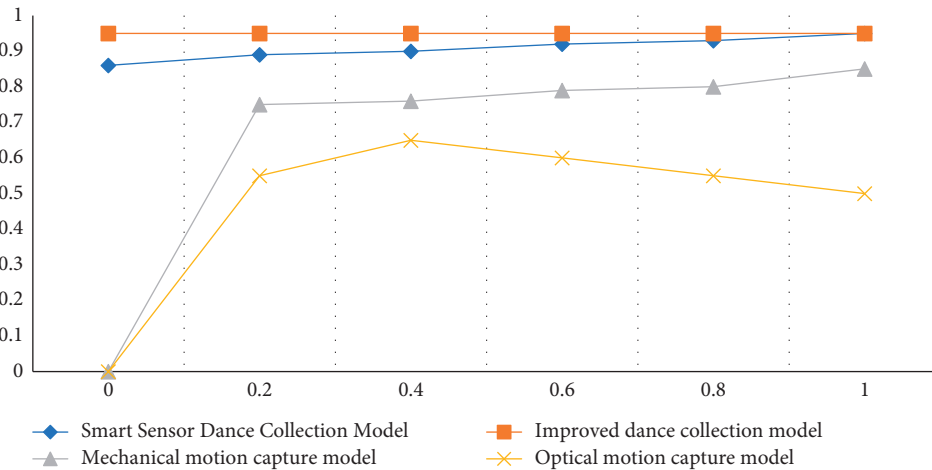


FIGURE 8: ROC curve under the mixed test set.

we can also see that the ROC curve values of the article and the improved model are very stable, 0.96. The ROC value of the mechanical motion capture model is low, the ROC curve of the optical motion capture model is more tortuous, and the ROC value is also low. The experimental results also show that the performance of the music and dance acquisition model of the smart sensor is the best.

According to the data in Table 8 and Figure 8, we can conclude that after the test set runs, the performance of the four models has decreased to a certain extent, but the smart sensor dance acquisition model proposed in the article has the lowest degree of decline, and the performance after the decline is still highest among the 4 models, with an accuracy of 90.24% before model improvement and 93.16% after improvement. According to the ROC curves of the four algorithms, we can also see that the ROC curve value of the improved system is very stable, no matter in the test set or the mixed test set, the ROC value has been kept at 0.95, the ROC value before the improvement is stable within the range of 0.85–0.95, the ROC curves of the mechanical motion capture model and the optical motion capture model are more tortuous, and the ROC values are also lower. The experimental results further illustrate that the model proposed in this paper has the best performance.

5. Conclusion

Relying on intelligent sensor technology, this paper studies the real-time acquisition method of music and dance gestures. The research idea of the article breaks through the

traditional research method and provides a new idea for dance acquisition method with information and intelligent technical means. The general idea of the article is to pre-process the collected dance images, including image grayscale and background detection, to extract the character features in the video images. Then use the principle of intelligent sensor to realize the recognition of dance movements. Although some good achievements have been made in dance collection, there are still some deficiencies in some aspects. In the process of collecting dance video data, only one perspective data can be collected, and multiple perspectives cannot be covered. The dance database of this article has a small amount of information and cannot meet the needs of a large number of dance trainers. In the future research work, we should start to solve these problems.

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 regarding this work.

References

- [1] V. E. Gonzalez-Sanchez, A. Zelechowska, and A. R. Jensenius, "Correspondences between music and involuntary human micromotion during standstill," *Frontiers in Psychology*, vol. 9, 2018.

- [2] S. A. Taylor, "Butoh: a bibliography of Japanese avant-garde dance," *Collection Building*, vol. 31, no. 1, pp. 15–18, 2012.
- [3] S. Saha, S. Ghosh, A. Konar, and A. K. Nagar, "Gesture recognition from Indian classical dance using Kinect sensor," in *Proceedings of the 2013 5th International Conference on Computational Intelligence, Communication Systems and Networks*, pp. 3–8, Madrid, Spain, June 2013.
- [4] S. Mozarkar and D. Warnekar, "Recognizing bharatnatyam mudra using principles of gesture recognition," *International Journal of Computer Science & Network*, vol. 02, no. 04, 2013.
- [5] M. S. Febrianti, E. Hidayat, A. I. Wuryandari, A. S. Prihatmanto, and C. Machbub, "Preliminary result on gesture recognition of sgeh penguten dance using hidden Markov model," in *Proceedings of the 2016 6th International Conference on System Engineering and Technology (ICSET)*, pp. 169–172, Tamil Nadu, India, May 2016.
- [6] S. Saha, L. Ghosh, A. Konar, and R. Janarthanan, "Fuzzy L membership function based hand gesture recognition for bharatanatyam dance," in *Proceedings of the 2013 5th International Conference and Computational Intelligence and Communication Networks*, pp. 331–335, Mathura, India, September 2013.
- [7] C. K. Raju, "SVD-based feature extraction from time-series motion data and its application to gesture recognition," *Computing Reviews*, vol. 56, no. 8, p. 511, 2015.
- [8] E. Jablonskis, "Adaptive gesture recognition system, transforming dance performance into music," *Jablonskis*, vol. 04, no. 14, pp. 21–39, 2009.
- [9] Y. Heryadi, M. I. Fanany, and A. M. Arymurthy, "Grammar of dance gesture from bali traditional dance," *International Journal of Computer Science Issues*, vol. 9, no. 6, pp. 23–42, 2012.
- [10] S. Boukir and F. Chenevière, "Design of a dance gesture recognition system," *Traitement du Signal*, vol. 3, no. 3, pp. 971–974, 2004.
- [11] A. Camurri, M. Ricchetti, and R. Trocca, "EyesWeb: toward gesture and affect recognition in interactive dance and music systems," *Computer Music Journal*, vol. 10, no. 23, pp. 18–29, 2000.
- [12] S. Boukir and F. Chenevière, "Compression and recognition of dance gestures using a deformable model," *Pattern Analysis & Applications*, vol. 7, no. 3, pp. 308–316, 2004.
- [13] N. Anbarsanti and A. S. Prihatmanto, "Dance learning and recognition system based on hidden Markov model. A case study: each traditional dance," in *Proceedings of the 2014 IEEE 4th International Conference on System Engineering and Technology (ICSET)*, pp. 1–6, Bandung, Indonesia, November 2014.
- [14] G. Chen and S. Li, "Network on chip for enterprise information management and integration in intelligent physical systems," *Enterprise Information Systems*, vol. 15, no. 7, pp. 935–950, 2021.
- [15] Q. Dong, Y. Wu, and Z. Hu, "Gesture recognition using quadratic curves," *Computer Vision - ACCV 2006*, vol. 3851, pp. 817–825, 2006.
- [16] D. Perzanowski, A. C. Schultz, W. Adams, E. Marsh, and M. Bugajska, "Building a multimodal human-robot interface," *IEEE Intelligent Systems*, vol. 16, no. 1, pp. 16–21, 2001.
- [17] J. Fritsch, M. Kleinhagenbrock, and S. Lang, "Multi-modal anchoring for human-robot interaction [J]," *Robotics and Autonomous Systems*, vol. 43, no. 2, pp. 133–147, 2003.
- [18] R. Stiefelwagen, C. Fugen, R. Gieselmann, H. Holzapfel, K. Nickel, and A. Waibel, "Natural human-robot interaction using speech, head pose and gestures," in *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2422–2427, Sendai, Japan, October 2004.
- [19] J. G. Han, J. Dalton, B. Vaughan, C. Oertel, C. Dougherty, and C. D. Looze, "Collecting multi-modal data of human-robot interaction," in *Proceedings of the 2011 2nd International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 1–4, Budapest, Hungary, July 2011.
- [20] Le. Sun, Q. Yu, D. Peng, S. Subramani, and X. Wang, "Fogmed: a fog-based framework for disease prognosis based medical sensor data streams," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 603–619, 2021.
- [21] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231–268, 2001.
- [22] T. B. Moeslund, A. Hilton, and V. Krieger, "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, no. 2, pp. 90–126, 2006.
- [23] R. Z. Hu, A. Hartfiel, J. Tung, A. Fakhri, J. Hoey, and P. Poupart, "3D Pose tracking of walker users' lower limb with a structured-light camera on a moving platform," in *Proceedings of the CVPR 2011 WORKSHOPS*, pp. 29–36, Colorado Springs, CO, USA, June 2011.
- [24] J. Shotton, R. Girshick, A. Fitzgibbon et al., "Efficient human pose estimation from single depth images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 12, pp. 2821–2840, 2013.
- [25] L. A. Schwarz, A. Mkhitarian, D. Mateus, and N. Navab, "Human skeleton tracking from depth data using geodesic distances and optical flow," *Image and Vision Computing*, vol. 30, no. 3, pp. 217–226, 2012.

Research Article

Exploring the Integration Model of Industry Chain Information System Based on Energy Internet and Its Key Technologies

Anxun Qi 

Shuifa Group Co LTD, Jinan 250101, China

Correspondence should be addressed to Anxun Qi; 20047682@sdufe.edu.cn

Received 11 April 2022; Revised 16 May 2022; Accepted 17 May 2022; Published 31 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Anxun Qi. 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.

Many enterprises have elevated industrial chain management to an important strategic factor for successful operation in a competitive market environment. There are many reasons for the low success rate of enterprise information construction, one of which is that in the process of information construction, only the research of information technology and the construction of information systems are emphasized, but not the integrated management of information resources, and the strategy, mode and method of information resource management are not studied. As a solution for sustainable energy development in the future, the energy Internet has become a hot topic of research in academia and industry today. Integrated innovation, at the enterprise level, is the key to improving core competitiveness and maintaining a sustainable competitive advantage. In order to achieve a deep integration of information technology and energy technology, the energy Internet must be supported by a corresponding system, i.e., an energy information system. In this paper, we propose an information system integration model based on the nature of the industry chain, describe its functional components, and analyze the key technologies of integration. The results show that the integrated model achieves the highest value of each output solution in the time period of 4h, and the calculated transmission capacity of the line is 0.7pu. The integrated model of the information system is optimal for the adaptation of the energy environment of each distributed power source in the energy Internet. Therefore, this study provides theoretical support and practical reference for further research on the construction and mechanism of the integrated energy Internet in China's theoretical and business communities, and provides important insights into the construction of the mechanism and countermeasures for the integration of the industrial chain in energy Internet enterprises.

1. Introduction

Today's world has shifted from the era of industrial economy to the era of information and knowledge-based service economy [1]. The environment facing modern enterprises has changed dramatically and is still changing [2]. The popular application of electronic data interchange technology, barcode technology, electronic fund transfer technology and Internet/Intranet technology has provided extensive and effective technical support for industrial chain management [3]. Industrial chain management is a systematic, integrated and agile management mode, which integrates new ideas and technologies of modern management today [4]. The modern business environment has prompted enterprises to integrate their resources with external resources to better respond to customer needs, create more value for customers, and create

more profits for enterprises [5]. Enterprises use information systems as a tool to develop and utilize a large amount of information resources [6]. Although the development of information technology provides advanced processing means and tools for the development and utilization of enterprise information resources, the effectiveness of its application mainly depends on the mode and method of enterprise information resource management [7].

Therefore, it is important to study the information management system to improve the quality of information management and to ensure the efficiency of the company's management and its ability to handle daily affairs. Information system is a complex network with vulnerability [8]. The vulnerability of an information system is caused by defects in the design and implementation of the hardware and software, related protocols and security strategies that

make up the system [9]. When information failures such as interruptions or delays occur in the energy information system, error messages cause the information system to react incorrectly and affect the normal operation of the physical grid [10]. In practice, the variety and scope of information system maintenance operations are expanding, and the number of O&M personnel is increasing, while the diversity and complexity of information systems place higher demands on the professionalism of O&M personnel [11]. Therefore, the energy Internet covering many microgrids can be regarded as a unified energy resource pool, and the orderly flow of energy among microgrids can be controlled through information interaction, which can effectively suppress the limitations of microgrids and ensure the safety and stability of the whole energy Internet.

The successful implementation of industrial chain management cannot be achieved without two enablers, organizational structure and information technology [12]. By adopting industrial chain management, enterprises can improve their market competitiveness [13]. Especially when global industrial chains are formed, the flexibility of enterprises to participate in global competition increases [14]. Industrial chains are not only required to eliminate obvious, inefficient, and non-value-creating activities, but also aim to streamline and intellectualize a series of value-creating activities and eliminate inter-process continuity through the integration of systems [15]. The core of integrated management emphasizes the use of integrated thinking and concepts to guide management practices, which is essentially the integration of elements and complementary advantages. Therefore, it is necessary to strengthen the internal integration and external network of the enterprise, to emphasize internally the parallel development of fully integrated functional departments, to participate in the production of knowledge and information from their respective perspectives, and to research and develop the use of expert systems and simulation models.

The innovation points of this paper are as follows:

- (1) Based on the existing theories of integrated innovation and energy Internet, the paper defines the concept, connotation and characteristics of enterprise integrated energy Internet from two levels based on the operation mechanism and key technologies of energy Internet, breaking through the previous scholars' definition of the concept of network.
- (2) Information resource integration as a management practice is one of the core tasks of enterprise informatization, and promoting information resource integration can help eliminate information silos and realize information resource sharing.
- (3) The integration theory is applied to the study of enterprise energy Internet, and the integration mechanism and model of enterprise integrated energy Internet are constructed from the perspective of integration of goals, information, resources, knowledge, culture and other elements.

The research framework of this paper contains five major parts, which are organized as follows:

The first part of this paper introduces the research background and significance, and then introduces the main work of this paper. The second part introduces the work related to the integration model with the information system of industrial chain, the energy Internet. In the third part, the method of building the functional structure of the industrial chain and the distributed process management model are sorted out, so that the readers of this paper can have a more comprehensive understanding of the method of building the industrial chain information system integration model based on the energy Internet. The fourth part is the core of the thesis, which completes the description of the operation mechanism and key technology analysis of the Energy Internet from two aspects: the operation mechanism and business model analysis and the key technology analysis of the Energy Internet. The last part of the thesis is the summary of the whole work.

2. Related Work

2.1. Industry Chain Information System Integration Mode.

The changes in the environment faced by modern enterprises, especially the development and application of e-commerce with the Internet as the core technology, are triggering a major change that is causing profound changes in the way people live and behave. At the level of the internal chain, management emphasizes the coordination of entities and activities within a single enterprise: the information system is built and implemented mainly using Intranet technology and MRPII/ERP management software. Integrated information systems must require strong compatibility, cross-platform and cross-language operation, and provide reliable access to Web applications. Therefore, enterprises now need to be able to organize, search, and access truly valuable information through the integrated management of information resources.

Yang et al. argue that integrated enterprise information resource management is the process of organizing and integrating information resources in the internal and external environment of an enterprise to support business change and organizational transformation, using the mission, purpose, and strengths of the enterprise as the integration point [16]. Hao and Wang analyze the product development strategy, process, and performance of the pharmaceutical industry in the context of globalization [17]. Sun et al. argue that information resource integration is divided into data integration, content integration, and process integration, and that process integration is based on data integration and content integration [18]. Griffiths proposed that the integration of R&D and marketing activities, as well as the integration of R&D activities and user requirements, has shown an important contribution to NPD under the conditions of technological and market uncertainty [19]. Sun et al. argue that information resource integration includes system integration and sharing, software integration and sharing related to business processes, and information integration and sharing [20].

Therefore, in the face of the challenges and requirements of the new era, it is of great academic value and practical

significance to explore the mode and method of enterprise information resource management from the new perspective of integration and to form a set of mode and method of enterprise information resource management to meet the new era background and requirements.

2.2. Energy Internet. At present, the contradiction between China's economic development and energy structure is becoming more and more prominent, and the demand for transformation of energy production and consumption patterns is becoming stronger. The information in the Energy Internet covers power supply, grid and load in all aspects, including static basic data describing system parameters, dynamic data monitoring system operation status, and various analysis data. In the context of the energy Internet, the traditional O&M model with equipment as the maintenance object is developing into a new model with users as the service object, which has changed from the current information and communication O&M objectives.

Li et al. introduced the definition, key technologies, and development models of the energy Internet [21]. Users are both consumers and producers. Wang et al. discussed the relationship between the Internet of Things, smart grid, and cloud computing with the energy Internet [22]. Ren et al. proposed the technical connotation and technical characteristics of the energy Internet, and analyzed the technical elements and technical forms of the energy Internet [23]. Hu et al. Most of the electrical energy in the energy Internet is generated randomly, and thus it is impossible to predict the availability of a certain moment [24]. Yan studied the architecture of the energy Internet and analyzed the key technologies in the process of developing the energy Internet [25]. Stable and reliable power supply is the result of the joint action of all power sources on the whole grid, and these characteristics match well with cloud computing characteristics.

In order to cope with the current severe information and communication operation and maintenance situation and continuously improve the information and communication business level, it is urgent to study a new model of information and communication system operation and maintenance management in the context of the energy Internet. On this basis, we focus on the integration mode of information system of industrial chain based on energy Internet, as well as the physical architecture and technical route of energy information system, and put forward the key technology of energy information system construction.

3. Construction Method of Industrial Chain Information System Integration Mode Based on Energy Internet

3.1. Method for Construct Functional Structure of Industrial Chain. Business activities of industrial chain management are divided into three levels: strategic level, tactical level and operational level. When conducting inter-enterprise information system integration, it is extremely important to

ensure the security of network information in the process of transmission [26]. Energy Internet operational mechanisms and business models are the core of the Energy Internet and are the key to improving comprehensive energy efficiency and building an integrated energy service system [27]. According to the main features of the energy Internet and combined with the Internet technical architecture system, the study proposes a three-layer physical-information-application construction model, which is shown in Figure 1.

First of all, the strategy level is the decision-making level, which is long-term and relatively stable, and determines the general policy, policy and overall planning and design of the enterprise chain, including decisions on positioning, production, inventory, sales and transportation. When a service requester needs to invoke a service it first goes to the service agent to find the service it needs and gets information on how to invoke the service, and then goes to invoke the service published by the service publisher based on this information. Defining the network topology model, the characteristic path length of the network is given by the following equation.

$$L = \sum_{i \neq j} \frac{2 * d_{ij}}{[N * (N - 1)]}, \quad (1)$$

i, j --Any two nodes in the network, N --Total number of nodes in the network, d_{ij} --The shortest distance between two nodes

It connects the FMIS system with other subsystems of the MIS and subsystems within the FMIS system into an organic whole by means of data interfaces. The size of the enterprise is determined in the marginal cost of organizing transactions within the enterprise is equal to the marginal cost of organizing the same transactions in the market or in another enterprise [28]. So that they support each other, call each other, have common possession of information and common access to data in order to achieve a holistic effect that cannot be achieved by a single application. The total number of neighboring nodes of a node is called the degree of the node. Looking at the model as a whole, at any point in the network, the degree distribution function of that point is :

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)} \propto 2m^2 k^{-3}, \quad (2)$$

m --Number of adjacent nodes, k --Parameters for calculating degrees

Therefore, it is necessary to streamline the scale of enterprises, reduce the degree of integration of enterprises, and externalize certain inefficient departments and functions from within the enterprise, leaving only the core functions of the enterprise with the most competitive advantage. The functional structure of the industrial chain is shown in Figure 2 below.

Secondly, the tactical and operational levels are execution levels that develop short- and medium-term plans and are characterized by short and dynamic role cycles. Service agents give services the mechanism to publish (register and classify) themselves and their services, and also provide

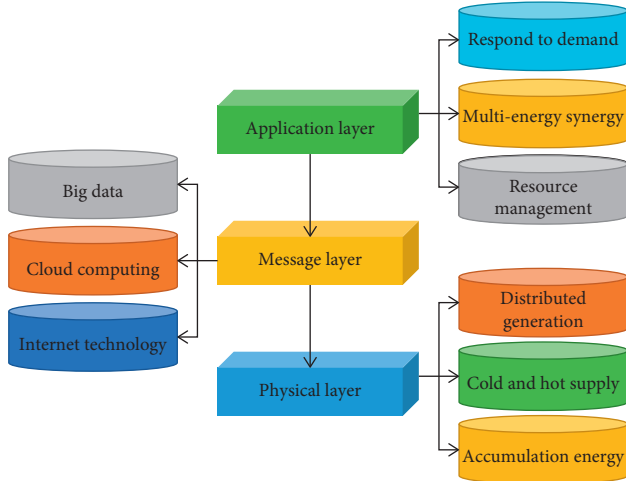


FIGURE 1: Schematic diagram of energy internet engineering construction architecture.

service requests with the mechanism to find the services they need [29]. The firm or cluster of firms that is in the position of an innovation agent is called a core firm, and a vertical network is formed between the core firm and the suppliers of upstream and downstream relationships as well as users. The aggregation coefficient of nodes is defined as the ratio of the number of edges E_i that actually exist between these k_i nodes to the maximum number of edges that theoretically exist, as:

$$C_i = \frac{2E_i}{k_i(k_i - 1)}. \quad (3)$$

Further, the aggregation factor of the network can be defined.

$$C = \frac{1}{N} \sum_{i=1}^N C_i. \quad (4)$$

According to the role mechanism of enterprise information resources, combined with the distribution, status and demand of information resources in the enterprise, we can generally consider enterprise information resources integration from three aspects: the goal of enterprise information resources integration, the content of integration and the scale of integration. Application Integration (EAI) middleware is deployed in each department and data center, and departmental data exchange platform and central data exchange platform are established respectively to realize data exchange and sharing among departments. To improve the information communication network within the system, realize the two-way flow and effective utilization of electrical information data within the power grid, and on this basis, realize the effective control of the power grid dispatching center on the distribution network. For each virtual node, j is a certain neighboring node that is making a decision. The transmission probability based on dynamic information is defined as

$$P_{i,j} = \frac{e^{-t_j}}{\sum_{l \in N(i)} e^{-t_l}}. \quad (5)$$

Finally, the operational level is related to daily business activities such as inventory management, production activities, equipment management, specific job scheduling, etc. In the process of information resource integration, enterprises must determine the objectives of integration according to their specific status and needs, so that the integration can be targeted and the role and potential of enterprise information resources can be fully and effectively played. Through rapid data exchange, real-time integration of data from relevant departmental business systems or interaction between business systems, so that leaders and departments can make comprehensive use of the results of information integration to assist in decision-making and management.

3.2. Distributed Process Management Model. The industrial chain system is based on the synergy and information sharing among multiple enterprises [30]. The industrial chain is not static, and as the market environment changes, the members in the chain will also change, with new members joining in and old members possibly dropping out. According to the main features of the energy Internet, combined with the Internet technology architecture system, the study proposes a three-layer construction model of physical-information-application. Because of this, many companies have elevated industry chain management to an important strategic factor for successful operation in a competitive market environment. The distributed process management model is shown in Figure 3 below.

First, the model includes interfaces for interoperability and communication between components in five categories. The physical layer focuses on addressing multiple energy sources interconnection and grid vertical intelligent interaction. The industry can design and build a dedicated network system according to their actual needs to best meet their needs. The diversity of information formats is manifested in structured data of business systems, semi-structured data of network systems and unstructured data of system log files. The distributed generation and exchange power from the main grid must be equal to the energy LAN consumption power, so the power balance equation is shown below :

$$\sum_i^{NDG} P_i(t) + P_{Grid}(t) = P_{Load}(t) + \sum_{n=1}^{NPHEV} P_{PHEV}(t). \quad (6)$$

$P_{Load}(t)$ --Load power, $P_{PHEV}(t)$ --Charge/discharge power

The global data that needs to be transmitted, exchanged, processed and shared throughout the MIS and among various subsystems for the purpose of information integration of commercial banks are mainly the main-frame business transaction data of the enterprise. Both the degree of node and the waiting time of stream have an impact on the routing efficiency of the system. Therefore, the weights of node degree and waiting time for calculating the probability of selecting adjacent nodes for transmission are adjusted by adjusting this parameter in the following way :

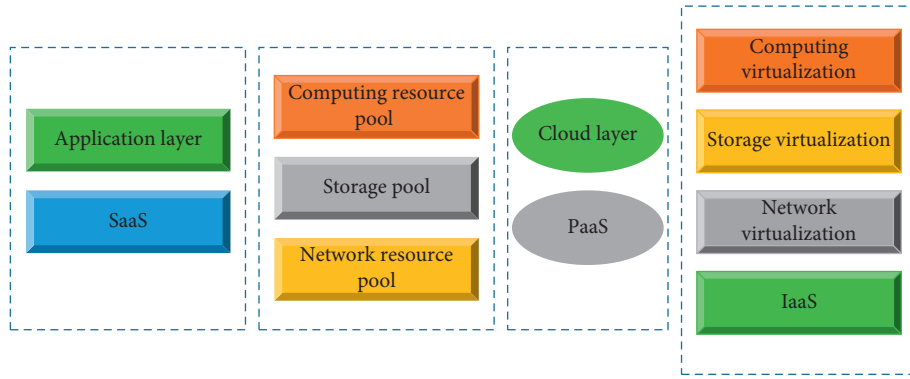


FIGURE 2: Functional structure of industrial chain.

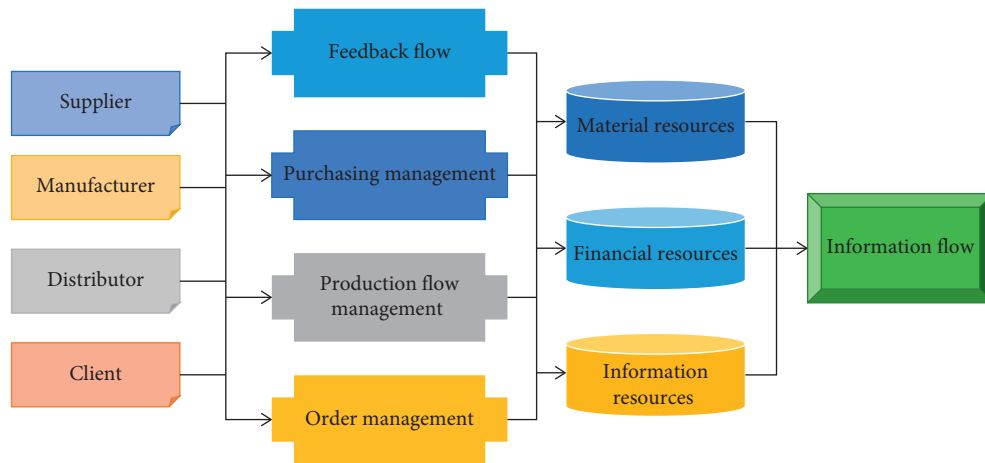


FIGURE 3: Distributed process management model.

$$P_{i,j} = \frac{k_j^\alpha e^{-(1-\alpha)t_j}}{\sum_{l \in N(i)} k_l^\alpha e^{-(1-\alpha)t_l}} \quad (7)$$

Second, the process definition tool interface, client interface, call application interface, enabling component interoperability interface, management and monitoring program interface. The information layer applies big data cloud platform technology to address the interconnection of massive information from multiple sources. The information can be divided into 3 categories according to the role of static information such as grid equipment parameters and topology, business information such as power generation and load demand changes and transactions, and dynamic control information for grid operation and monitoring. By using Internet network for integration, enterprises can build an integrated industry chain information system with good scalability and interoperability (most network systems are based on TCP/IP protocol) without building their own private networks or paying expensive leased line fees. The Bessel formula is then used to calculate the standard error of the original data and replace it with the replacement value, and the formula for calculating the width of the replacement value code is:

$$B = \frac{V}{(F \times 2^\xi)} \quad (8)$$

V --Optional range of data acquisition, F --Optional gain of data acquisition, ξ --Resolution of collector

The clustering formula is as follows The clustering formula is:

$$E = \sum_{j=1}^K \sum_{x \in C} |x_j - \bar{x}_j|^2 \quad (9)$$

E --Information clustering function, x_j --Data set, \bar{x}_j --Mean data set, C --Cluster center of data set, K --Cluster collection

Local data is the data exchanged and shared within each subsystem of the MIS system, and transferred, exchanged and manipulated within the subsystem. It inherits and develops the advantages of various systems such as design resource management, design process management and information management, and applies parallel engineering methodology, network technology, database and object-oriented technology. For network actors, the network provides various formal and informal communication channels, thus promoting the flow of knowledge, which is very important for the dissemination of new ideas and

innovation. In order to simplify the operations in the application calculations, the output power can be calculated by the following equation :

$$P_{pv} = P_{STC} \frac{G(T)}{G_{STC}} (1 + k(T_c - T)_r). \quad (10)$$

Finally, each interface develops a series of APIs to standardize the function invocation and data exchange formats between the two sides of the component. The application layer focuses on solving the new energy production and consumption, market and trading model based on Internetization. In addition to having the traditional grid operation and management data, the energy Internet, due to the access of new energy sources, for wind and photovoltaic energy sources, must be based on a large amount of meteorological background data for the prediction of electrical energy output. Develop a software architecture that integrates the various information and application resources involved in the business process of an enterprise into one information system, so that different information and applications can be shared and interacted with each other and operate as an integrated whole. In the process of data management, each subsystem handles and processes private data to produce local data, and through the exchange and processing of local data, global data is generated within the subsystem for delivery and exchange to the whole system. Its emergence provides a new solution to the problem of enterprise information integration.

4. Analysis of Operation Mechanism and Key Technologies Based on Energy Internet

4.1. Analysis of Operation Mechanism and Business Model. The industrial chain is a concept containing four dimensions: value chain, enterprise chain, supply and demand chain and spatial chain. It is an up-and-down and dynamic chain intermediate organization formed by enterprises in the same industry or different industries with the goal of satisfying users' demands based on specific logical links and spatial and temporal layouts. A suitable operation mechanism and business model can provide energy enterprises, power sales companies, users, owners of distributed resources and other types of subjects to participate in the implementation path of energy Internet operation and create an open energy Internet ecology.

First, the price-based multi-energy cooperative operation mechanism forms a method for optimizing the operation of cross-region multi-energy systems at the city level. In order to accurately reflect the characteristics of the regional integrated energy network, the network of energy production, transmission and use constituted by multiple energy forms needs to be properly modeled. The characteristics of intermittency, instability and small scale affect the safe and stable operation of the microgrid. the solution process of 2 intelligent optimization algorithms for energy internet energy economic adaptation optimal scheduling is shown in Figure 4.

Compared with PSO, LSMFO has a faster convergence rate in the early stage and a weaker search capability in the later stage, while the optimization performance presented at the end is also less than ideal. The deep integration of multi-party industry chain is a prerequisite for the establishment and normal operation of energy Internet enterprises. The information and communication operation and maintenance services are controlled in real time to improve service quality and ensure safe and reliable business. The weak link of the whole energy information system is no longer limited to the internal physical system, and the failure of the information link will become more common and diverse, which will most likely further endanger the energy system. The system completes the cloudization of information network and hardware resources, and provides computing capability, storage capability, data analysis capability, and application integration capability for the application system layer, etc. The distribution of DE and PSO in the energy Internet energy economic adaptation regarding battery decoupling is shown in Figure 5.

Secondly, the virtual power plant day-ahead scheduling model based on double-layer planning can realize dynamic prediction of distributed new energy generation capacity in a specific region. The establishment and development of energy Internet enterprises must break the barriers between industrial chains and achieve the deep integration of all industrial chains. The real-time dynamic characteristics of information are mainly reflected in two aspects: first, the ability to obtain system operation status data and user demand information in real time; second, the ability to realize high-performance computing of massive data. The connection between various enterprises in the industrial chain needs to be closer, and the resource allocation in all aspects needs to be more flexible and reasonable. In the Internet local routing mechanism, the algorithm based on dynamic information is calculated based on the length of the node's own data queue and the queue length of neighboring nodes, and the T and t pairs under different α values are shown in Figure 6 below.

To carry out hierarchical management of the generation and maintenance enterprises, determine a pool of competent and trustworthy generation and maintenance enterprises, and establish a long-term cooperation mechanism. The enterprises in the chain are divided into basic and auxiliary enterprise members, so as to further analyze the key processes and auxiliary processes in the chain and their characteristics. After the power layer converter gets the command value, it starts to adjust the power flow, and the results are shown in Table 1.

Finally, drawing on the Internet thinking model, an open integrated energy operation service platform should be built to provide data capabilities and business capabilities for each subject to carry out integrated energy operations. It should also model the information network composed of multiple communication technologies and media, reflecting the mutual influence and interaction between the energy network and the information network. It is to model the value chain system of the industry chain, determine the starting position of the industry chain, describe the role and role of

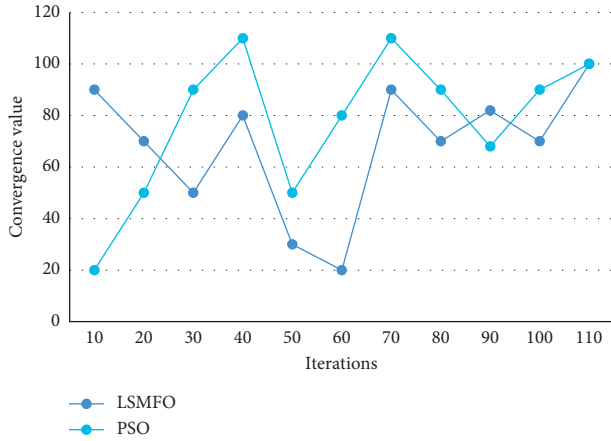


FIGURE 4: Convergence characteristics of two algorithms.

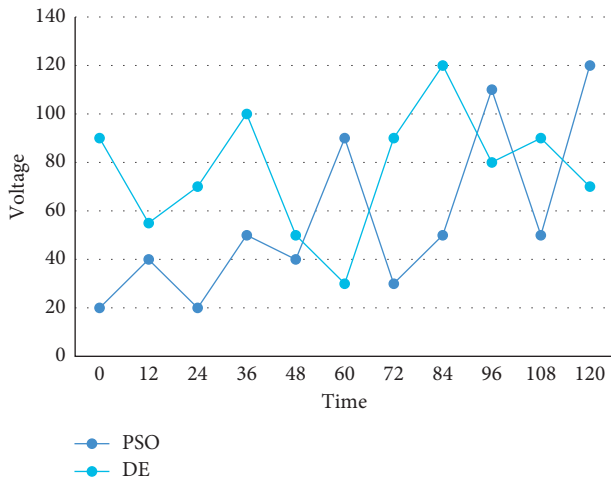


FIGURE 5: Comparison of a set of solutions of DE and PSO on battery.

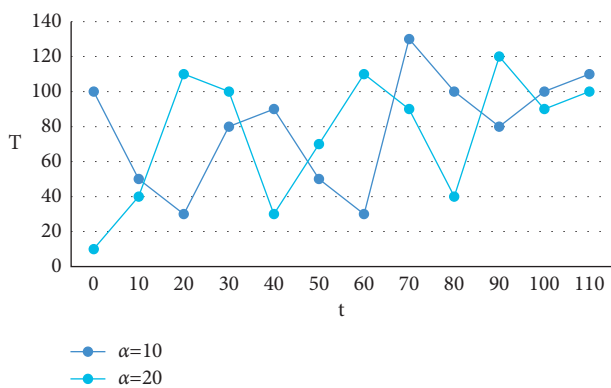


FIGURE 6: T and t with different alpha values.

TABLE 1: Active current distribution of each line after using the power routing strategy.

	Active power/Pu	Cost
Line parameters	0.5	6
MAS calculated value	0.4	3
Active control value	0.7	9

enterprises in the industry chain, and analyze the problems and their root causes in the value system of the industry chain. Due to the development of enterprises in many aspects such as massive renewable energy access, large-scale energy storage and information technology demand interaction, a large amount of data information from multiple sources such as energy production, energy storage information and user consumption characteristics will be generated. Therefore, it is necessary to complete the control of information and communication operation and maintenance from the organizational structure, and establish an internal management organization that can fit the scope of operation and maintenance business and adapt to market changes. The increasing integration of information link and energy system is no longer a simple information flow regulating energy flow, but a real integration of information flow and energy flow.

4.2. Analysis of Key Technologies Based on Energy Internet. The traditional work-based organizational approach and organizational management model, with its emphasis on work as the core and individuals as subordinate to work, has neglected the critical resources and important components of organizational competitiveness of today's organizations. Complex network theory provides another theoretical basis for mechanism-based modeling and simulation of inter-continental energy networks. The data transmission rate of the communication line is shown in Figure 7.

In order to realize the prerequisites provided by the integration of information and energy systems to the construction and operation of energy Internet enterprises, and to truly realize the linkage of supply and demand among multiple subjects, it is necessary to break the boundary restrictions between different industries and strengthen the information security of all parties. The energy Internet-based industry chain information system aims to build a convenient and efficient information collection, processing, analysis and display platform with the high concurrency, high reliability, high scalability and low cost of cloud computing. Its key technology is aimed at supporting the physical energy grid to achieve integrated energy inter-connection, efficient and flexible distribution network, and intelligent and efficient network.

Firstly, the third generation semiconductor Si C is used to form a wide forbidden band power device, which has lower switching loss and lower on-state voltage drop to enhance the conversion efficiency of the device. Since the power is generated centrally and the grid is dispatched and managed in a unified manner, the safety and stability of the system can be ensured by realizing effective dispatch management of the grid. The processing of these public data is realized through the functional modules of FMIS system to realize the processing, transmission and sharing of data among the functional modules. Therefore, it is not the management of each specific database by PDM, but the management of information integration on a macro level, the rise of local automation to overall automation, and the transformation of local optimum to overall optimum. Under

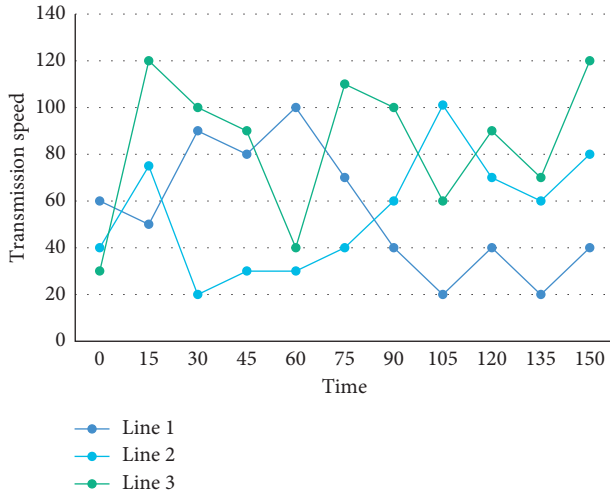


FIGURE 7: Data transmission rate of communication line.

ideal communication conditions, there is no delay in data transmission, and the node voltage can immediately respond to load changes, while with the communication delay, Figure 8 is a schematic diagram of the dynamic change of voltage at node 600 under different communication delay conditions.

Secondly, the multi-port DC circuit breaker topology with shared transfer branches and the development and application of three-port DC circuit breakers greatly reduces the number of electronic devices and reduces the volume and cost. The Energy Internet requires data centers to provide richer information services and a high-performance distributed computing environment for data mining and assisted decision-making. It integrates the right business data, business specifications, business procedures, policies and controls. Security services to control access between industry chain member companies, information is passed between industry chain member companies, security is a key issue that needs to be addressed and must ensure that information is only available to the appropriate members in an accurate manner. the openness of the PDM kernel is mainly reflected in the system object-oriented modeling technology to build the system management model and information model, and provide object management mechanisms to achieve product information Management Table 2 shows a set of optimal results of the integrated model for energy adaptation and scheduling of each distributed power source in the energy Internet based on environmental costs.

When the time period is 4h, each output solution reaches the highest value, and the transmission capacity of the line is calculated to be 0.7pu The integration mode of the information system is optimal for the energy environment adaptation of each distributed power source in the energy Internet.

Finally, the engineering and technology scheme of control and protection integration and compact design is proposed to realize the functions of coordinated control of multi-terminal flexible DC distribution network, one-key system start-stop, and seamless switching of multiple

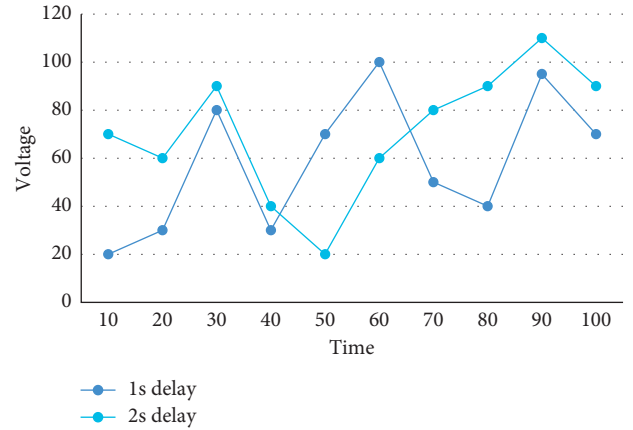


FIGURE 8: A schematic diagram of the voltage dynamic change of node 600 under different communication delay conditions.

TABLE 2: Output solutions of integrated mode in distributed generation of energy Internet.

Time period (h)	1	2	3	4
Battery (power)	-25	-18	14	28
Fuel cell (power)	3.672	4.663	5.763	6.871
Microcomputer (power)	8.54	9.65	10.26	11.76

operation modes. It can combine the business processes and business logic of an enterprise to provide predetermined functions to assemble the application system that suits the needs of the enterprise. The key used for encryption is different from the key used for decryption, and the decryption key cannot be calculated from the encryption key. The responsibility of the service provider is to invoke its own information system according to the w SDL document provided by the service proxy center, implement the corresponding functions, and register and publish them in the proxy center so that they can be accessed by other applications, accept and handle calls from Web service requestors.

5. Conclusions

Enterprise informatization is the process of reshaping enterprise competitiveness through informatization, in which the integrated management of information resources is the key to enterprise informatization. The security framework of integrated industrial chain information system is a topic between the industrial chain field and the network security field, which is of great relevance today when the network security issue is of great concern. The energy Internet is more complex than the smart grid currently being carried out, and places higher demands on network transmission performance, data storage management, data analysis and mining, and information interaction. From the depth of integration and utilization of enterprise information resources, it will gradually develop from information integration to knowledge integration, and then to human integration, so as to realize the "human-centered" management idea in the enterprise. According to the direction and interrelationship of

enterprise integration, the integration methods of enterprise energy Internet can be summarized as industry-production network, industry-academia-research network, government-industry-academia-research network and its government-industry-academia-research-gold network. Carrying out research on the theoretical basis and practical path of the energy Internet and realizing the interconnection of energy resources are of strategic importance to China's future energy security. Therefore, this paper proposes an industrial chain information system integration model based on the energy Internet, starting from the research framework of enterprise information system integration, from the perspective of the overall construction of the entire industrial chain management system, and analyzing its characteristics and key technologies. The energy Internet and the industrial chain interact with each other, and jointly promote the interactive integration of enterprise integrated energy Internet, and make the network produce synergistic operation effect, information sharing effect, common vision effect, complementary capability effect and knowledge transfer effect, and promote the development of cooperative relationship among members of integrated energy Internet.

Data Availability

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

Conflicts of Interest

The authors declare no conflicts of interest

Acknowledgments

This work in this article was supported by 2019 Shandong key R & D Program

References

- [1] Z. Yan, Z. Tao, W. Rui, and Y. Liu, "A model predictive control based distributed coordination of multi-microgrids in energy internet," *Acta Automatica Sinica*, vol. 43, no. 008, pp. 1443–1456, 2017.
- [2] X. Zhu, "Research on key technologies and applications of energy internet blockchain," *E3S Web of Conferences*, vol. 118, no. 19, Article ID 01003, 2019.
- [3] Q. Qu, X. Sun, S. Xu, and L. Qiang, "Research on renewable energy access to power grid based on energy internet technology," *IOP Conference Series: Earth and Environmental Science*, vol. 558, no. 5, p. 6, Article ID 052063, 2020.
- [4] Y. Fu, X. Yang, P. Yang, Y. Wong, and S. Zheng, "Energy-efficient offloading and resource allocation for mobile edge computing enabled mission-critical internet-of-things systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 1, pp. 1–16, 2021.
- [5] H. Cui, C. Wang, and H. Liu, "Monitoring and analysis of distributed new energy resources based on the internet of things," *IOP Conference Series: Earth and Environmental Science*, vol. 714, no. 4, p. 8, Article ID 042025, 2021.
- [6] K. Wang, J. Yu, Y. Yu, and Q. Yirou, "A survey on energy internet: architecture, approach, and emerging technologies," *IEEE Systems Journal*, vol. 12, pp. 2403–2416, 2017.
- [7] Q. Sun, Y. Zhang, H. He, M. Dazhong, and Z. Huaguang, "A novel energy function-based stability evaluation and non-linear control approach for energy internet," *IEEE Transactions on Smart Grid*, vol. 8, pp. 1195–1210, 2017.
- [8] K. Wang, X. Hu, H. Li, L. Peng, and G. Song, "A survey on energy internet communications for sustainability," *IEEE Transactions on Sustainable Computing*, vol. 2, no. 3, pp. 231–254, 2017.
- [9] M. Zeng, J. Cheng, Y. Wang et al., "Primarily research for multi module cooperative autonomous mode of energy internet under blockchain framework," *Proceedings of the CSEE*, vol. 37, no. 13, pp. 3672–3681, 2017.
- [10] A. S. Sani, D. Yuan, J. Jin, and G. Longxiang, "Cyber security framework for internet of things-based energy internet," *Future Generation Computer Systems*, vol. 93, no. APR, pp. 849–859, 2018.
- [11] M. Rakhmatullaev and U. Karimov, "Models OF integration OF information systems IN higher education institutions," *Society Integration Education Proceedings of the International Scientific Conference*, vol. 5, p. 420, 2018.
- [12] Y. Wang and L. Zhang, "Design of computer information system integration based on metadata," *Journal of Physics: Conference Series*, vol. 1992, no. 2, p. 5, Article ID 022085, 2021.
- [13] F. Xia and P. Wang, "Research on power line communication technology of energy internet based on blockchain," *IOP Conference Series: Earth and Environmental Science*, vol. 714, no. 4, p. 8, Article ID 042085, 2021.
- [14] C. Feng and X. Liao, "An overview of "energy + internet" in China," *Journal of Cleaner Production*, vol. 258, no. S1, Article ID 120630, 2020.
- [15] M. S. Asif and H. Gill, "Blockchain technology and green supply chain management (GSCM) – improving environmental and energy performance in multi-echelon supply chains," *IOP Conference Series: Earth and Environmental Science*, vol. 952, no. 1, p. 10, Article ID 012006, 2022.
- [16] L. Yang, Y. Zou, M. Xu, X. Yicheng, and Y. Dongxiao, "Distributed consensus for blockchains in internet-of-things networks," *Tsinghua Science and Technology*, vol. 27, no. 5, pp. 817–831, 2022.
- [17] W. J. Hao and X. M. Wang, "A study on ecological operation of web IP industrial chain," *Ecological Economy*, no. 03, pp. 49–55, 2017.
- [18] Q. Sun, R. Wang, and D. Ma, "An islanding control strategy research of we-energy in energy internet," *Proceedings of the Chinese Society of Electrical Engineering*, vol. 37, no. 11, pp. 3087–3098, 2017.
- [19] A. Griffiths, "The impact of industrial internet of things on supply chains," *Supply chain brain*, vol. 23, no. 1, p. 13, 2019.
- [20] Q. Sun, M. Zhao, Y. Chen, and D. Ma, "Optimal energy flow of multiple energy systems in energy internet," *Zhongguo Dianji Gongcheng Xuebao/Proceedings of the Chinese Society of Electrical Engineering*, vol. 37, no. 6, pp. 1590–1598, 2017.
- [21] C. Li, X. Li, S. Tian et al., "Challenges and prospects of risk transmission in deep fusion of electric power and information for energy internet," *Automation of Electric Power Systems*, vol. 41, no. 11, pp. 17–25, 2017.
- [22] J. Wang, L. Wang, and J. Lin, "Energy internet morphology and its technical support system," *Dianli Zidonghua Shebei/Electric Power Automation Equipment*, vol. 37, no. 4, pp. 1–10, 2017.

- [23] G. Ren, M. Wu, and M. Yu, "Automatic recognition algorithm of information architecture reliability based on energy internet network topology," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 2706646, 9 pages, 2021.
- [24] J. Hu, Y. Chen, X. Ren, and Y. Yixian, "Blockchain-enhanced fair and efficient energy trading in industrial internet of things," *Mobile Information Systems*, vol. 2021, Article ID 7397926, 13 pages, 2021.
- [25] R. Yan, "Optimization approach for increasing revenue of perishable product supply chain with the Internet of Things," *Industrial Management & Data Systems*, vol. 117, no. 4, 2017.
- [26] Z. Shi, J. Jiang, and S. Jing, "Application of industrial internet identifier in optical fiber industrial chain," *ZTE Technology: English version*, vol. 18, no. 1, p. 7, 2020.
- [27] J. Chi, Y. Li, J. Huang, and L. Jing, "A secure and efficient data sharing scheme based on blockchain in industrial Internet of Things," *Journal of Network and Computer Applications*, vol. 167, Article ID 102710, 2020.
- [28] Q. Li and L. Chen, "Internet + industrial chain integrating poverty reduction and old-age care with online education," *People's forum · academic frontier*, vol. 000, no. 013, pp. 92–95, 2018.
- [29] Y. Sun and J. T. LiLi, "Research on the application of big data in regional industrial supply chain," *Journal of Physics: Conference Series*, vol. 1883, no. 1, p. 4, Article ID 012167, 2021.
- [30] Z. Shen, "Integrated supply chain design models: a survey and future research directions," *Journal of Industrial and Management Optimization*, vol. 3, no. 1, pp. 1–27, 2017.

Research Article

Comprehensive Evaluation Method of Digital Economy Development Level Based on Complex Network Model

Xingjun Shi 

Zhejiang Industry and Trade Vocational College, Wenzhou, Zhejiang 325002, China

Correspondence should be addressed to Xingjun Shi; sxj@zjitc.edu.cn

Received 8 April 2022; Revised 15 May 2022; Accepted 18 May 2022; Published 31 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xingjun Shi. 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 economic globalization and artificial intelligence modern society, the rapid economic growth mainly relies on the vigorous development of digital information technology. Simultaneously, digital technology promotes economic revolution and produces a digital economy, which has become an important concept of China's economic development. In order to more scientifically and accurately describe the development status of my country's industrial structure and digital economy industry, this paper takes each industry of the national economy as the research object and puts forward the idea of using complex network theory to study the relationship between input and output. Based on the complex network theory, this paper analyzes the interrelationships between various industries in the complex system. By referring to the relevant data of the national economic input and output in recent years, the input-output correlation network model between the various industries in the national economic system is established. Using the method of combining statistical analysis and comparative analysis, quantitative analysis and qualitative analysis, carry out detailed and in-depth statistical analysis from the input-output correlation network model and model attributes so as to point out the status and role of various industries in the national economy and so as to analyze and study the industrial structure of the national economy.

1. Introduction

With the development of the emerging economy, the digital economy has gradually entered people's field of vision and has received more and more attention. The scale of my country's digital economy continues to expand, and the innovative Internet, such as big data technology, information technology, and cloud computing, has further expanded the reach and service scope of the digital economy, becoming a key engine for promoting economic growth and driving high-quality economic development. The exchange of different elements between industries drives the interaction between various industries. The industrial sector of the national economy forms an industrial network model through the input-output relationship. The characteristics of the industrial network structure affect the evolution of the industrial structure. The emergence of the digital economy industry depends on the full development of the digital economy. The construction of its indicator system is based

on traditional products and services, but it also covers the digital Internet economy and the advantages of the digital economy that it brings [1, 2].

On this basis, as so to better explore the development of the digital economy, this paper first combines the input-output correlation analysis with the complex social network model and then builds an input-output correlation model based on complex network theory so as to systematically explore the relationship and position of each industry from the overall level. The main network attributes of each industry, such as intensity distribution, agglomeration coefficient, etc., are analyzed, and the degree and impact of digital economy industries on other industries are discussed. Then, by referring to the compilation modus of the traditional financial inclusion index, combined with the current characteristics of my country's digital economy, using the analytic hierarchy process based on the coefficient of variation to determine the index weight, and establish an evaluation system [3]. On the basis of the overall index, the

digital economy index is subdivided into three different dimensions: coverage breadth, depth of use, and degree of digital support services, which measures the development level of the digital economy from two macro- and micro-levels and puts forward corresponding conclusions. The conclusions presented give reasonable suggestions.

2. Definition of Digital Economy Industry and Related Research Methods

This paper conducts a social network analysis on the characteristic structure and degree of correlation between the digital economy industry and other industries, which is not only a supplement and improvement of the theory in the digital economy field but also a new exploration of applying social network analysis methods to the digital economy industry. At the beginning of the whole, each industry is regarded as the research object, combined with the input-output modus in the national economy, the direct consumption coefficient is used as the weight to weigh the edges connected to each industry, and the strong correlation is extracted by setting the critical value, and then the strong correlation is constructed. An industrial complex network graph of relationships: using indicators such as network density, centrality, agglomeration coefficient, and block model, this paper analyzes the characteristics of my country's industrial development; by extracting digital economy-related industries, measuring the internal connection between digital economy industries and other industries, and exploring my country's overall and local perspectives. The characteristics of the industrial network also reveal the characteristics of the input-output correlation network of the industrial structure of my country's national economy [4–6]. Applying the complex network to the study of the industrial structure network is a new attempt in the economic system, and there are still many works. It can be further explored, and the above content also provides ideas for the development research of other industries. The research on enriching the digital economy industry and other industrial structures also has certain theoretical significance. This paper first studies the definition standards and development of the digital economy at home and abroad and then sets out from the actual situation in my country, breaks through the traditional perspective, from the perspective of the innovative digital economy, and scientifically and comprehensively establishes an indicator system that can summarize the digital economy from a multidimensional perspective. It is not only helpful to promote the deepening of digital economy research but also a useful supplement to the development of the traditional economy.

2.1. Research Status at Home and Abroad. Jaakkola et al. used the S-shaped curve to describe the spillover process of high-tech enterprises such as information and communication technology [7]. Some scholars explored the production efficiency of different industries in the United States and found that industrial productivity improvement is related to information and communication technology. It is relatively

obvious that information and communication technology has a pulling effect on its related industries [8, 9], and there is a strong spillover effect between the information and communication technology industry and other industries, and it is the key engine for promoting economic growth [10], simultaneously, the information and communication technology manufacturing industry shows high heterogeneity, indicating that the information and communication manufacturing industry is more widely used in the industry. There are also scholars who hold different views on the role of the digital economy. Bart pointed out that although the digital economy has developed rapidly, it remains to be discussed whether it has played its greatest role [11]. Through further quantitative calculation, Chattopadhyay concluded that if the banking system lacks a digital economy, the national economy will lose 1% [12]. Demircuc and Klapper stated in the World Bank report that digital finance could help the poor to obtain deposits or borrowings at a fixed and reasonable ratio [13] and build a safer payment transaction environment by establishing a credit system. Indian economist Sarma draws on the construction modus of the United Nations Human Development Index (HDI) [14] by setting the availability and specific usage of financial services and banking penetration as the main indicators, using Euclidean distance and linear efficacy function to correlate. The combined modus reasonably measures the basic status of the digital economy in different countries [15–17].

According to the definition of the digital economy, the China Academy of Information and Communications Technology and Liu divide the digital economy into two parts: industrial digitization and digital industrialization [18], of which the industrial digitization part refers to the integration of other industries and digital technologies [6, 19]. An application can increase and improve the output and efficiency brought by the use of digital technology and digital products in other nondigital industrial sectors of the national economy; digital industrialization includes digital product production and digital technology innovation, mainly covering electronic information manufacturing, information communication industry, Internet industry, and software service industry. To explore the information and communication technology industry in-depth, Wen used the social network model to measure the degree of relatedness of the information technology industry [20–22] and innovation co-operation space and found that the local spatial organization innovation characteristics of the cluster innovation network society are obvious, and the innovation subjects in the cluster area are more inclined to cooperate with the innovation subjects in other cluster areas; the latter adopts the input-output model to obtain a new generation of information technology. The main relevant departments of the industrial chain are used to measure the degree of correlation between the industrial chains by using the average sweeping steps method [23, 24].

The research of domestic and foreign scholars on industrial correlation and the digital economy industry provides an important reference for this article. As a new economic form arising from the development of the information technology revolution, most scholars study the

digital economy industry or application at the theoretical level. The regression method conducts an empirical analysis of the digital divide in different regions, while the research on the development of the digital economy industry and its degree of correlation with other industries is relatively rare, not in-depth enough, and lacks multilevel research. The research on the social network in foreign countries is earlier than that in China. A lot of explorations have been carried out in both theoretical and empirical aspects, and a series of rich research upshots have been obtained, providing important references and inspiration for the research of domestic industrial social networks.

2.2. Possible Innovations and Deficiencies of Existing Research

- (1) Innovation in research perspective. The digital economy industry is an emerging industry, so this paper takes the digital economy as a research perspective, which is an innovation. Drawing on the definition of the digital economy industry in other countries and considering the development status of my country's digital economy industry, the digital economy industry is defined as a digital technology-based industry. The sum of various economic activities is carried out to determine the main areas of my country's digital economy industry. This article compiles the index based on a large amount of microdata provided by Ant Financial Group and the relevant data from the statistical yearbook that can be found. The index not only covers the emerging digital economy industry but also integrates traditional economic businesses, such as banks. From the perspective of measuring the development status of the digital economy, it is reliable and representative. Not only does it provide a basis for government decision-making, but local governments can also use this index to judge the level of digital economic development in their region and learn from the development experience of other regions. Secondly, it provides a basis for decision-making reference for enterprises to make business decisions. Finally, the construction of a digital economy development index can also help to more accurately understand the development of the entire industry and increase the understanding of the digital economy.
- (2) Innovative research modus. Apply social network analysis to industrial structure analysis. The input-output method is widely used in the study of industrial structure, but it cannot visualize the closeness and concentration trend of interindustry relations and has certain limitations, and the social network analysis modus can make up for the shortcomings. Therefore, based on the input-output analysis method, this paper introduces the perspective of network analysis, which can dig deeper into the correlation of various industries and provide support for identifying the development

status of the industrial structure. Through the construction of my country's industry association network, the role of my country's industry is analyzed, and it is divided into different sections to study its function and role. Simultaneously, taking the industrial system as a whole reveals the overall characteristics of my country's industrial development; simultaneously, the various industrial elements that make up the complex industrial network model also reveal the level and structure of the entire network and clearly explain the relationship between industries from the overall and local levels and the tightness of connections and the relationship between industry behavior across the network. The logarithmic efficacy function is used to carry out the dimensionless treatment of the index so as to maintain the stability of the index and alleviate the influence of extreme values. Finally, according to the characteristics of the digital economy index, the weighted arithmetic average modus is used to construct the digital economy index.

- (3) Due to the relative limitations of the data, the research on the current situation of industrial development in this paper only analyzes the year 2015, which cannot fully demonstrate the actual situation of the current industrial development. In addition, due to the limited academic level of individuals, there is also a lack of in-depth research on the mechanism and impact of industry associations and research on the prediction of industry development. In addition, due to the availability of data and the compatibility of data from different institutions, the research angle and indicators may be slightly single, and there is no in-depth research and exploration from a certain dimension.

3. Analysis of Network Characteristics of China's Digital Economy Industry

The input-output method proposed by Leontief is the most commonly used modus in industrial linkage research. However, the input-output method can only reflect the relationship between industrial departments and cannot reflect the structural characteristics of industrial departments in detail, and the input-output analysis modus needs to meet basic assumptions such as homogeneity and proportionality. Therefore, based on the input-output table, this paper uses the social network analysis method to establish an industrial network model. It is possible to analyze further the relationship between the digital economy industry and other industries, as well as the overall structural characteristics of my country's industrial associations so as to provide support for the development of the digital economy industry. As shown in Figure 1, the digital economy plays an irreplaceable role in promoting the economic development of all walks of life, deepening economic reform, and promoting industrial upgrading by using mathematical thinking and digital technology and platforms.

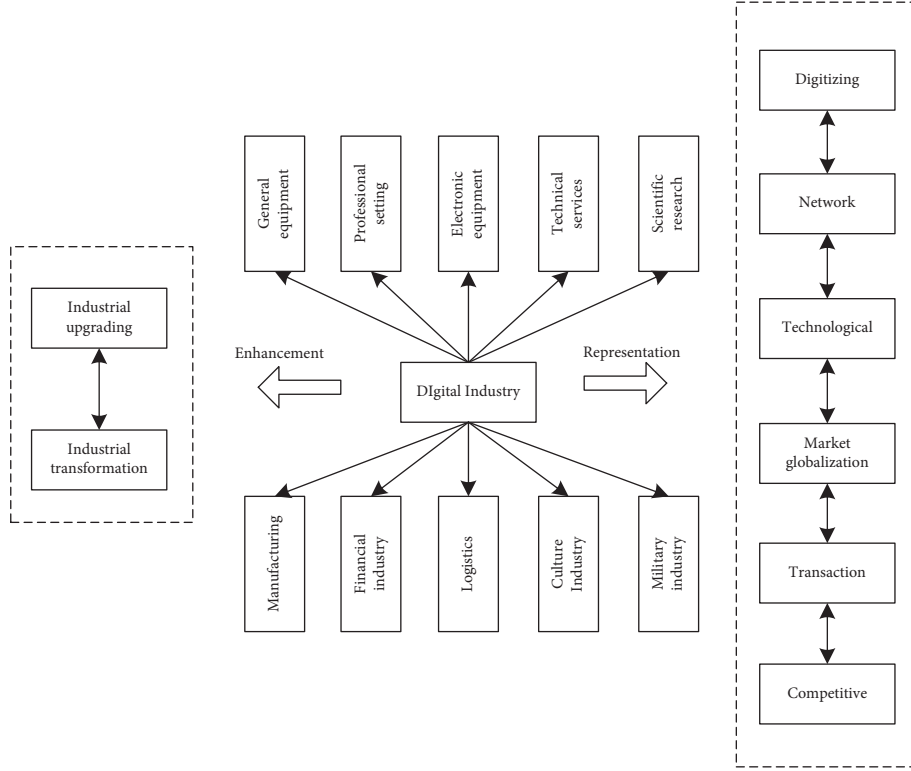


FIGURE 1: Digital industry structure chart.

3.1. Industrial Network Model Construction. The paper adopts the idea of input share to connect the edges of the strong correlation relationship with the direct consumption coefficient of the industry greater than the average value, weights the edges with the direct consumption coefficient, and establishes the industrial network model $N=(V, E, W)$, where V is the node industry, and E is the directed edge between nodes. It reflects the correlation between industries, and $W=\{W_{ij}\}$ is the weight, which represents the direct consumption coefficient of industry j to industry i .

- (1) Determination of Nodes According to my country's inter-regional input-output table, my country is divided into 42 major departments. Therefore, the inter-regional industrial space network constructed in this paper has a total of 42 nodes.
- (2) Determination of side and side rights, the input-output relationship between industries in each region is regarded as the edge of the network. The direct interregional industry correlation coefficient is introduced as the edge. The interregional industry direct correlation coefficient is introduced as the weight of the edge, and the critical value is set to determine the existence of the edge. Calculate the direct correlation coefficient equation (1) below. Among them, X_j is the output of j industry, X_{ij} , and it is the intermediate input of i industry to j industry. When constructing the industrial network model, the direction of input and output is considered. Therefore, for each node, there are actually four edges connected to it, node 1 is connected to four

directed edges, and each of the edges has a different weight.

$$a_{ij} = \frac{x_{ij}}{x_j} \quad (i, j = 1, 2, \dots, n). \quad (1)$$

(3) Theil index method

Theil index modus was first proposed by Theil in 1967. Its biggest advantage is that it can decompose the gaps between regions and within regions, which is more conducive to exploring regional development gaps and causes, and formulating more targeted policy. Therefore, it is more authoritative and convincing to select the Theil index to study the development gap of the national digital economy. The calculation method of the Theil index of the overall difference in the development of the national digital economy is shown in the following formula (2):

$$T = T_b + T_w = \sum_{k=1}^k del_k \log \left(\frac{del_k}{(n_k/n)} \right) + \sum_{k=1}^k del_k \left(\sum_{i \in g} \frac{del_i}{del} \log \frac{(del_i/del_k)}{(1/n_k)} \right). \quad (2)$$

T_b and T_w represent the digital economic development gap between regions and the digital economic development gap within the region, respectively, n represents the number of individuals, K represents K groups, and each group is g_k ($k = 1, 2, \dots, K$), the number of individuals in the K group g_k

is n_k , and del_i and del_k represent the digital economic development level of individual i and the overall digital economic development level of group K , respectively.

3.2. Analysis of the Overall Characteristics of the Industrial Network. The comprehensive index of the digital economy development level is the basic data for the study of regional differences in the development of the digital economy. The paper mainly starts from the three dimensions of infrastructure, digital transaction, and digital industry and establishes a comprehensive evaluation index system for the development level of the digital economy, as shown in Table 1. The comprehensive index modulus measures the development level of my country's digital economy. The larger the obtained comprehensive development index of the digital economy, the higher the development level of digital economy, and vice versa, the lower the development level of the digital economy.

The paper uses the Theil index method to calculate the development level of the national digital economy. Overall differences, examining the differences in the development of digital economy between regions and within regions in my country's current situation, the upshots calculated by Theil index method are shown in Table 2. The upshots, respectively, draw a map of the regional differences in the digital economy and the internal differences of the seven regions. Figure 1 shows a map of the differences in the digital economy.

It can be seen from Figure 2 that since 2013, the overall regional differences in the development level of my country's digital economy have shown a W-shaped trend. The overall difference showed an expanding trend, rising from 0.3014 in 2014 to 0.7770. Specifically, from 2013 to 2014, the overall regional difference in the development level of my country's digital economy showed a downward trend, from 0.7032 to 0.3014, and increased significantly from 2014 to 2016, from 0.7032 to 0.3014. 0.3014 rises to 0.8354, which decreased from 2016 to 2017, from 0.8354 to 0.7215, and then began to rise again, reaching 0.7770 in 2018, showing an overall upward trend. The difference in digital economy development between regions is basically similar to the overall regional difference in the level of digital economy development and has a strong positive correlation with the overall regional difference in the level of digital economy development. Differences and differences in digital economic development between regions have the opposite trend, showing a downward trend as a whole, from 0.1779 in 2013 to 0.1538 in 2018, but the change is small. Specifically, from 2013 to 2015, the difference in the development of the digital economy within the region showed an upward trend, rising from 0.1779 in 2013 to 0.1918 in 2015, and began to decline slightly from 2016, and then began to decline overall.

Judging from the differences in the development of the digital economy within the seven regions of my country, at the beginning of the whole, the difference in the development of the digital economy within the North China region is the largest. Although there has been a certain decline in recent years, it is still the highest among all regions. Second,

the development of the digital economy in all regions of my country has shown a downward trend in recent years. Third, the differences in the development of digital economy in Northeast and North China are basically the same, and the development of digital economy in Northwest and Southwest regions shows a consistent trend. The impact of intra-regional differences and inter-regional differences was on overall differences and the impact on the development of the digital economy. This paper calculates the contribution rate of regional differences and intraregional differences to the overall differences in the development of the digital economy, as shown in Table 3.

4. Comprehensive Evaluation of Digital Economy Development in Various Regions of China

To further explore the reasons for the regional differences in the development of my country's digital economy, to put forward more effective policy suggestions for the current situation of the regional differences in the development of my country's digital economy, the paper uses the grey correlation modulus to discuss the main reasons that affect the development of my country's digital economy, to narrow the country's digital economy and it shows the regional differences in the development of the digital economy and more effective suggestions for realizing the coordinated development of the national digital economy. According to the actual situation of my country's digital economy development, the paper believes that the reasons for the regional differences in the development of the digital economy can be mainly classified into three aspects: digital foundation, digital transaction, and digital industry. According to the calculation upshots of the grey correlation model, the following conclusions are drawn: Among them, the digital foundation has the highest degree of relevance, and its relevance ranks first, followed by digital transactions and digital industries, showing a ranking structure of digital foundation > digital transactions > digital industries.

Specifically, at the beginning of the whole, among all the influencing factors, the digital foundation has the greatest impact on the differences in the development of the national digital economy and is the main reason for regional differences. In recent years, with the development of the digital economy, various regions have accelerated the construction of digital economic infrastructure. However, due to the different levels of economic development in various regions, there will be a large gap in the investment in digital economy infrastructure, resulting in a large gap in the construction level of digital economy infrastructure in various regions, and infrastructure is the basis for the development of the digital economy. As a result, the development gap of the digital economy in various regions has shown an increasingly obvious trend. The second is the impact of digital transactions and digital industries on the development gap of the regional digital economy. Both digital transactions and digital industries are based on the Internet, breaking through geographical restrictions to a certain extent. Still, simultaneously, the economic development level of each

TABLE 1: Comprehensive index system of digital economy development level.

Primary indicator	Secondary indicator	Measurement method	Unit
Base installation	Cable length	Cable length per square kilometer	kilometer
	Computer users	Every 100 households use a computer	Platform
	Internet access volume	Internet broadband access users per 10,000 people	Household
Digital transaction	Internet usage	per 10,000 mobile internet users	Household
	The integration of digital economic entities	The number of enterprises with e-commerce transactions per 10,000 friends	Individual
	E-commerce sales	E-commerce sales per 10,000 people	Billion
Digital industry	E-commerce purchases	E-commerce purchases per 10,000 people	Billion
	Digital industry investment	Fixed asset investment per 10,000 people	Billion
	Digital industry human resources	Proportion of employees in information transmission services	%
	Digital industry turnover	Information service income per 10,000 people	Million

TABLE 2: Calculation results of the Theil index for the development level of the digital economy in the interval.

Years	All difference index	Intraregional difference index								Regional difference index
		North	Northeast	East	Central	South	Southwest	Northwest	Intraregional	
2013	0.703	0.449	0.050	0.273	0.075	0.165	0.036	0.031	0.177	0.525
2014	0.301	0.433	0.196	0.231	0.195	0.117	0.028	0.043	0.188	0.113
2015	0.757	0.506	0.075	0.247	0.048	0.183	0.076	0.062	0.191	0.565
2016	0.835	0.496	0.029	0.134	0.033	0.161	0.044	0.061	0.152	0.682
2017	0.721	0.461	0.012	0.176	0.029	0.166	0.065	0.041	0.153	0.567
2018	0.777	0.467	0.026	0.185	0.025	0.148	0.069	0.037	0.158	0.623

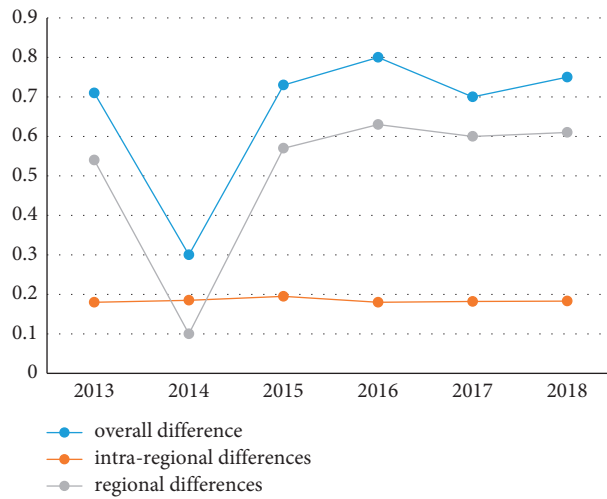


FIGURE 2: Map of regional differences in the digital economy.

TABLE 3: The contribution rate of regional differences and intraregional differences to the overall differences in the digital economy development.

Years	Intraregional contribution rate	Contribution rate between regions
2013	0.253001901	0.746998099
2014	0.624425624	0.375574376
2015	0.253257645	0.746742355
2016	0.182713242	0.817286758
2017	0.212899485	0.787100515
2018	0.197951304	0.802048696

region will also affect digital transactions and digital industries. Therefore, the two are important to the regional differences in the development of the national digital economy.

5. Conclusion and Recommendation

According to the previous research, the new technological revolution centered on the new generation of information technology has promoted the digital economy; it will give birth to many new industries. The development of the digital economy will expand the scope of production possibilities and accelerate industrial transformation and upgrade and promote high-quality economic development. Digital technology spawns new innovation models and infuses technological innovation with new impetus. This shows that taking advantage of the strategic opportunities brought by digital transformation accelerates the update of innovative ideas and it is particularly important to foster digital transformation, develop new R&D institutions, and optimize the policy environment for innovation.

5.1. Summary

- (1) The communication equipment, transportation equipment, computer, and other electronic equipment industries are in the central position. This may be because the tertiary industry, such as the transportation industry itself, has the function of linking other industries. This industry advantage is inherently unique. Therefore, as so to develop the regional economy, the government should pay attention to the close connection between industries. However, the internal structure of the tertiary industry has obvious hierarchical differentiation. Among them, the wholesale, retail industry, financial industry, transportation, warehousing, and postal industry have high centrality and occupy a dominant position in the entire industrial system. In contrast, the communication equipment, computer, and other electronic equipment manufacturing industries are ranked lower in emerging service industries such as high-tech industries, and almost all the resources of these industrial subgroups are used for internal consumption, only for internal industrial development. Provide resource support, resulting in a relatively unbalanced industrial structure.
- (2) The digital economy has realized the possibility of rapid economic development for economically backward regions and also laid the foundation for some middle-aged job seekers to obtain digital economic services with wider coverage and greater use depth. It is imperative to become high-tech intellectual talents. Physical talents in traditional industries will gradually be eliminated by society. In addition, the practice of the digital economy in recent years also shows that digital mobile payment technology provides a relatively strong ability to reach customers. With the advent of the era of

intelligence, the integration of the digital economy into people's lives will only become deeper and more common, sharing high-quality resources for education and medical assistance in different regions. For remote areas, increasing the coverage of the digital economy and reducing the cost of use are the top priorities for everyone; and if the digital economy is limited to developed areas, it will not only lead to a widening gap between the rich and the poor but also lead to the occurrence of unbalanced social structure in the long run.

5.2. Prospect

- (1) According to the previous research, industries with high digital economic strength are better at obtaining innovative resources from other industries to further expand the strength of the industry but neglect to build their own influence radiation range. As an agglomeration industry of digital economic resources, it should exert its own influence from two aspects. At the beginning of the whole, within the subgroup, give full play to the influence of the core industry of the digital economy and export resources to the original subgroup members to improve the subgroup. The overall strength of the group's internal network enhances the stability of the industrial network structure and builds a foundation for the digital economic relationship network. Simultaneously, it is necessary to give full play to the role of the intermediary of the sub-groups, enhance the output of the superior innovation resources of the subgroups to the sub-groups, actively explore more available innovation revenue resources in the industry, and strive to expand the sub-groups while improving their own innovation strength. The scale of the group, such as the transportation industry, can be further combined with the digital economy to develop services such as robot home delivery or intelligent rail transit so that the core industry can play its own advantages while driving the integrated development of other service industries.
- (2) The industrial age before the digital age emphasized economic scale; that is, the production cost of industry decreased with the increase of a single production scale, and the cost reduction could only depend on the increase of production quantity, which is a kind of decentralized development. Economic model: at present, the digital economy pays more attention to the development of a coordinated economy. In other words, under the condition of open resources, by strengthening the overall planning between different organizations, group development can bring greater economic benefits than self-development. The different characteristics of different regions in my country, such as industrial structure, innovation resource distribution, and digital resource endowment, also determine the

unique locational environmental advantages, digital resources, scientific, technological innovation resources, and other advantages of cities such as Beijing and Shanghai, which make these cities have advantage point for an internationally competitive digital economy.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The authors thank the Youth Foundation of Social Science and Humanity, China Ministry of Education, Research on Construction of Rural e-Commerce Development Index: Based on the Survey of 100 Typical Taobao Towns (no. 18YJC790135).

References

- [1] Y. Z. Zhang, "The development ideas and main tasks of the digital economy driving the industrial structure to the mid-to-high end," *Economic aspect*, vol. 34, no. 3, pp. 86–88, 2018.
- [2] X. S. Zhao, "Research on the transformation and upgrading of Chinese manufacturing driven by the digital economy," *Economic Theory and Practice*, vol. 18, no. 12, pp. 36–38, 2017.
- [3] P. F. Zhuo, *Research on the Transformation and Upgrading of China's Industrial Structure Promoted by Informatization*, Beijing University of Posts and Telecommunications, Beijing, China, 2017.
- [4] C. Yu, C. F. Wang, and W. Y. Zhuang, "Research on the evaluation of the new generation of information technology industry based on the perspective of dynamic evolution," *The Information of the Science*, vol. 36, no. 5, pp. 110–113, 2018.
- [5] J. H. Xing and H. Q. Gong, "Research on the evaluation of independent innovation capability of my country's new generation of information technology industry," *East China Economic Management*, vol. 31, no. 3, pp. 100–104, 2017.
- [6] R. H. Li and R. Huang, "Research on my country's digital industrialization and industrial digitalization mode," *Technology Management Research*, vol. 39, no. 16, pp. 129–134, 2019.
- [7] H. Jaakkola, H. Tenhunen, and A. Latvala, "Information technology and changes in industry," *European Journal of Information Systems*, vol. 1, no. 1, pp. 3–11, 1991.
- [8] K. Stiroh, "Are ICT spillovers driving the new economy," *Review of Income and Wealth*, vol. 5, no. 1, pp. 30–33, 2002.
- [9] P. S. Heo and D. H. Lee, "Evolution of the linkage structure of ICT industry and its role in the economic system: the case of Korea," *Information Technology for Development*, vol. 1, no. 3, pp. 1–31, 2018.
- [10] R. F. Aroche and S. G. Muniz, "Modeling economic structures from a qualitative input-output perspective," *Munich Personal RePEc Archive*, vol. 25, no. 2, pp. 292–296, 2012.
- [11] V. Bart, "Productivity and digitilization in europe:paving the road to faster growth," *Digi-World Economic Journal*, vol. 4, no. 100, pp. 107–115, 2015.
- [12] S. Chattopadhyay, "Financial inclusion in India: a case-study of West Bengal," *Reserve Bank of India Working Paper*, vol. 32, no. 5, pp. 1087–1095, 2011.
- [13] K. Demircuc and L. Klapper, "Measuring financial inclusion: the global finindex database," *Policy Research Working Paper Series*, vol. 4, no. 5, pp. 25–30, 2012.
- [14] M. Sarma, "Index of Financial Inclusion-A measure of financial sector inclusiveness," *Berlin Working Papers on Money*, vol. 4, no. 7, pp. 34–36, 2012.
- [15] D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [16] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.
- [17] T. Beck, A. Demircuc-Kunt, and M. S. Martinez Peria, "Reaching out: access to and use of banking services across countries," *Journal of Financial Economics*, vol. 85, no. 1, pp. 234–266, 2007.
- [18] J. Liu, *Lecture Notes on Holistic Network Analysis: UCINET Software Practical Guide*, Gezhi Publishing House, Shanghai, China, 2009.
- [19] C. Zhou, X. Z. Cao, and G. Zeng, "The cluster network mode and evolution path of China's electronic information industry innovation," *Geographical Studies*, vol. 38, no. 9, pp. 2212–2225, 2019.
- [20] J. Wen, Z. J. Yan, and Y. Cheng, "The digital economy and the improvement of regional innovation capabilities," *Exploration of Economic Issues*, vol. 11, pp. 112–124, 2019.
- [21] H. L. Lu, Y. Zhou, and J. B. Tang, "Research on industry strong association network based on complex network," *Journal of Beijing University of Posts and Telecommunications*, vol. 16, no. 4, pp. 46–54, 2014.
- [22] K. Wan, Y. B. Liu, and X. J. Huang, "Research on inter-regional coordination of high-tech manufacturing industry chain based on input-output model," *Operations Research and Management*, vol. 28, no. 5, pp. 190–199, 2019.
- [23] Q. Sun, M. R. Han, and J. X. Lei, "Research on regional innovation association network based on social network analysis," *Research Management*, vol. 39, no. 12, pp. 78–85, 2018.
- [24] J. D. Luo, *Social Network Analysis Handout*, Social Science Literature Press, New York, 2010.

Research Article

Construction of Remote Sensing Model of Fresh Corn Biomass Based on Neural Network

Jianjian Chen,¹ Hui Zhang,² Yunlong Bian,³ Xiangnan Li,¹ and Guihua Lv¹ 

¹Institute of Maize and Featured Upland Crops, Zhejiang Academy of Agricultural Sciences, Dongyang, Zhejiang 322100, China

²Zhejiang Agricultural Technology Extension Center, Hangzhou, Zhejiang 310000, China

³Jiangsu Key Laboratory of Crop Genetics and Physiology,
Co-Innovation Center for Modern Production Technology of Grain Crops,
Key Laboratory of Plant Functional Genomics of the Ministry of Education, Yangzhou University, Yangzhou,
Jiangsu 225009, China

Correspondence should be addressed to Guihua Lv; lvgh@zaas.ac.cn

Received 1 April 2022; Revised 12 April 2022; Accepted 13 April 2022; Published 31 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Jianjian 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.

Corn has a high yield and is widely used. Therefore, developing corn production and accurately estimating corn biomass yield are of great significance to improving people's lives, developing rural economy and climate issues. In this paper, a 3-layer BP neural network model is constructed by using the LM algorithm as the training algorithm of the corn biomass BP network model. From the three aspects of elevation, slope, and aspect, combined with the BP neural network model of corn biomass, the spatial distribution of corn biomass in the study area is analyzed. The results showed that the average biomass per unit area of maize increased with the increase in altitude below 1000 m. There are relatively more human activities in low altitude areas, which are more active in forestry production. The best planting altitude of corn is 0 ~ 1000 m. When the altitude is higher than 1000 m, the corn biomass gradually decreases. In terms of slope, if the slope is lower than 15°, the biomass of maize increases with the increase in slope. If the slope is lower than 15°, the biomass of maize decreases gradually with the increase in slope. The biomass of maize on sunny slope was higher than that on shady slope.

1. Introduction

Corn is an important food crop. Corn has strong adaptability, wide distribution, multiple uses, and great potential for increasing production. It ranks third in the world in terms of sown area and total output after rice and wheat and is developing rapidly. As an important food crop, corn is also rich in nutrients [1–3]. Since corn resources are extremely abundant, cheap, and easy to obtain, they also have many biological activities, such as antioxidation, antitumor, hypoglycemic, improving immunity, and bacteriostasis, which have broad development and application prospects. Therefore, developing corn production and accurately estimating corn biomass yield are of great significance to improving people's lives, developing rural economy, and solving climate issues [4–6].

Remote sensing information model method has become one of the most important methods in biomass research. This method estimates forest biomass through the correlation between chlorophyll content information reflected in remote sensing images and forest biomass. Therefore, the correlation mathematical model between remote sensing spectral information and measured biomass was established, and the forest biomass was estimated by inversion [7]. The remote sensing information model method is not only suitable for biomass estimation in small areas but also shows good results in biomass estimation in large areas. According to the different mechanisms of using remote sensing to estimate forest biomass, it can be divided into multiple regression analysis fitting relationship method, artificial neural network method, and so on [8–10].

The correlation analysis between the vegetation index and crop physiological and biochemical parameters established by the reflectance of different bands of the crop canopy can build a remote sensing biomass estimation model under different conditions and realize the remote sensing estimation of crop biomass yield [11, 12]. The current sensors for constructing such characteristic bands or vegetation indices are multispectral [13], visible RGB camera [14], and hyperspectral [15]. The relationship models established by using canopy spectral information mainly include linear (univariate and multivariate) regression models, nonlinear (exponential, logarithmic) regression models, powerful machine learning regression algorithms, and deep learning algorithms [16, 17]. Because there is a strong relationship between plant biomass and its own physical structure information (height, leaf area, bulk density, etc.) Therefore, biomass estimation models can be established through canopy structure parameters, such as tree allometric models [18] and regression models [19].

The measurement of crop canopy structure parameters relies on detecting the true location of the crop canopy and the terrain below the canopy, which means that a digital surface model that accurately represents the terrain is required to derive the vegetation canopy structure [20]. At present, the main sensors used to obtain the three-dimensional structure information of crop canopy based on remote sensing platform are visible light RGB camera and lidar LiDARo RGB camera. Compared with other sensors, the cost is lower and the time resolution is higher. The consumer-grade UAV RGB system can obtain high-precision images, construct crop surface models and digital elevation models, and generate orthophotos, point cloud information, and so on to determine crop structure information through image processing [21] to establish a data basis for accurate estimation of crop biomass [7, 22, 23].

Bendig et al. [24] used an airborne RGB camera to estimate the average canopy height and then estimated the fresh biomass ($R^2 = 0.81$) and dry biomass ($R^2 = 0.82$) of barley through the established exponential model. Ballesteros et al. [25] used RGB camera-based estimates of vegetation cover, plant height, and canopy volume to evaluate dry leaf biomass and dry bulb biomass of onion with R^2 of 0.76 and 0.95, respectively. Liu et al. [26] estimated the aboveground biomass based on the visible light spectral information and texture features of the potato canopy at different heights. The results show that adding texture features can greatly improve the damage estimation accuracy. LiDAR has high spatial resolution, is not affected by ambient light, and is highly repeatable, so it can provide more accurate three-dimensional structure information of vegetation canopy at the field scale than RGB cameras [27–31].

The advantage of BP neural network lies in its strong fitting ability, especially in biomass remote sensing inversion. The BP model can use the brightness values of multiple spectral bands of multispectral data (such as TM, ETM+, and so on). The underlying regularity between forest biomass and its spectral reflectance properties can be accurately estimated [32]. The research by Guo et al. [33] shows that the

BP model shows superior characteristics in remote sensing inversion of forest biomass. Among them, Guo et al. [33] established a biomass remote sensing model based on neural network technology using Xiaoxing'anling TM images combined with ground survey data, with high accuracy. Their research also used conventional methods to establish regression models and compared the two methods. The analysis points out that the mechanism between biomass and remote sensing factors in the regression model of the conventional method is easy to elucidate, while the estimation accuracy of the neural network method model is higher [12–15, 17].

Based on the BP neural network platform, a nonlinear model system of forest biomass remote sensing is established. The relative errors of this model system in estimating the biomass of coniferous forest, broad-leaved forest, and mixed coniferous and broad-leaved forest are -1.147% , 2.38% , and 3.156% , respectively. The estimation accuracy is high. And we use the model system to generate the quantitative distribution map of biomass in the study area. The overall accuracy reaches 88%. The use of the geographic information system (GIS) and remote sensing images combined with ground information to estimate biomass has been able to estimate the amount of mass in a large area. Artificial neural network has good nonlinear approximation ability and plays an important role in the establishment of multidimensional nonlinear models. It can establish nonlinear remote sensing models with high fitting accuracy, making the types of remote sensing models more abundant. Wang et al. [34] established a nonlinear remote sensing model system based on a neural network platform and used this model to estimate the biomass of coniferous forest, broad-leaved forest, and mixed coniferous and broad-leaved forest in the study area [20, 32, 35, 36]. The overall estimation accuracy reached 88%. In this paper, based on neural network, the biomass remote sensing model of fresh corn is constructed and analyzed.

2. Methods and Theory

2.1. Artificial Neural Network. Artificial neural network is a new theoretical method that has emerged in recent years. It is a type of model established by simulating the human brain nervous system from the microscopic structure and function. It is a way to simulate human intelligence. Compared with other methods, it has the following characteristics: (1) does not require preassumption, only needs to learn sample training; (2) can be well suited to the advantage of noisy data. This kind of network depends on the complexity of the system and achieves the purpose of processing information by adjusting the interconnected relationship between a large number of internal nodes. Nonlinear adaptive information processing system consists of a large number of interconnected processing units. It is proposed on the basis of modern neuroscientific research results and attempts to process information by simulating the neural network of the brain to process and memorize information. The artificial neural network has the ability of self-learning and self-adaptation. It can analyze and grasp

the potential rules between the two through a batch of corresponding input-output data provided in advance and finally use the new input data to calculate the output according to these rules.

Artificial neural network is a nonprogrammed, adaptive, brain-style information processing. Its essence is to obtain a parallel and distributed information processing function through the transformation and dynamic behavior of the network and to imitate human beings to different degrees and levels. It is an interdisciplinary subject involving neuroscience, thinking science, artificial intelligence, computer science, and other fields and has a wide range of applications in various industries.

2.2. Model Construction

2.2.1. Selection of Independent Variables. The brightness value of remote sensing images fully reflects the spectral characteristics of ground objects. The visible light band (Band 1–3), near-infrared band (Band 4), and short-wave infrared band (Band 5, 7) of ETM+ are useful in vegetation identification and forest growth prediction. The derived bands of ETM+ images, namely, vegetation index, are also commonly used data in plant classification and identification. From the perspective of ecology and geology, elevation, slope, and aspect also affect the growth of plants. Therefore, the selected modeling independent variables include 3 categories: ETM+ original bands (Band 1–5, 7), derived bands of ETM+ data (vegetation index DVI, normalized vegetation index NDVI, and ratio index RVI), and geoscience information (DEM, SLOPE, and ASPECT).

The vegetation index is based on the different reflection characteristics of different vegetation to red light and near-infrared light, and the image operation of these two bands is carried out, so as to enhance the vegetation information and weaken the useless information, which has achieved the purpose of vegetation identification and plant biomass prediction. The calculation methods of each vegetation index are as follows:

Difference vegetation index is as follows:

$$DVI = DN_{NIR} - DN_R. \quad (1)$$

DN_{NIR} and DN_R represent the brightness values of ETM+ near-infrared band and red light band, respectively.

Ratio vegetation index is as follows:

$$RVI = \frac{DN_{NIR}}{DN_R}. \quad (2)$$

Normalized vegetation index is as follows:

$$NDVI = \frac{(DN_{NIR} - DN_R)}{(DN_{NIR} + DN_R)}. \quad (3)$$

The slope and aspect are obtained by performing terrain analysis on DEM in ArcGIS 9.3 software. In addition, the location data of various places are obtained in the field through GPS. Due to the influence of terrain and other factors, the location data will have certain errors. In order

to eliminate the errors caused by GPS positioning data, a neighborhood analysis of 3×3 pixels is performed on each independent variable; that is, the value of each pixel is replaced by the average value of 9 surrounding pixels (including itself). There is a great correlation between vegetation index, original zone, altitude, slope direction, and original zone. Therefore, principal component analysis was performed on 12 variables, and the autocorrelation between the variables was eliminated by dimensionality reduction and finally compressed into 8 principal components, which were used as the input variables of the model. Similarly, from the correlation between biomass and their respective variables, it can be seen that the linear correlation between biomass and each factor is not high, indicating that the relationship between biomass and remote sensing factors and geoscience factors should not be explained by a linear relationship, which should be explained by the BP neural network model with strong nonlinear fitting ability.

2.2.2. Construction of the BP Model. We construct a 3-layer BP neural network model (i.e., an input layer, a hidden layer, and an output layer), in which the number of nodes in the input layer is 8, the number of nodes in the hidden layer is 100, and the number of nodes in the output layer is 1. In this paper, the LM algorithm is used as the training algorithm of the forest biomass BP network model. The remote sensing factor data, terrain factor data, and biomass data of 40 plots are used as the training samples of the model, and 8 principal components obtained by principal component analysis of 12 original variables are used as model input variables. However, due to the dimension of each factor, the size and range of the data are different. The original data are normalized and then used for modeling to prevent the small data information from being overwhelmed by the large data information. The final simulation result is restored by denormalization processing. The biomass of the plot was normalized as the output variable of the model. According to the characteristics of the data, the hyperbolic tangent function \tanh is selected as the transfer function from the input layer to the hidden layer, and its value range is $(-1, 1)$, so as to ensure that the input value of the hidden layer is between $(-1, 1)$. The linear function purelin is used as the transfer function from the hidden layer to the output layer. The structure of the BP model is shown intuitively in Figure 1. In the figure, W is the weight matrix and b is the threshold matrix.

3. Results and Discussion

3.1. Evaluation of the BP Model. Figure 2 shows the change of the mean square error during the training process of the BP model. When epoch = 7, it reaches the preset minimum error of 0.001 (the error here is the normalized error term). Figure 3 is the fitting effect diagram of the BP neural network model. The predicted value is highly correlated with the measured value, indicating that the fitting effect of the model is ideal.

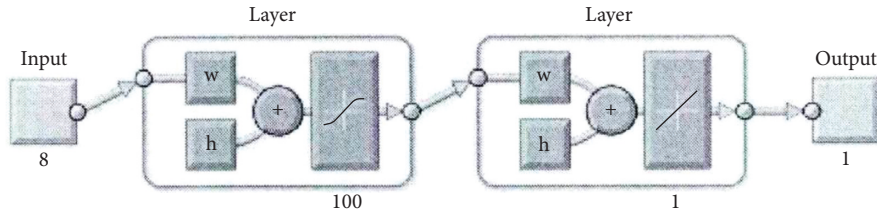


FIGURE 1: Diagram of BP network model structure.

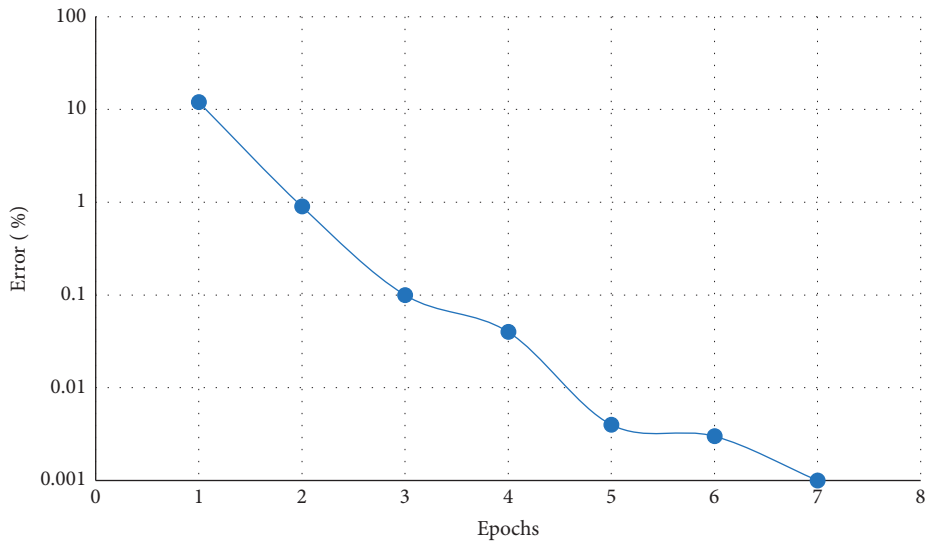


FIGURE 2: Variation of training error of the fresh corn biomass network model.

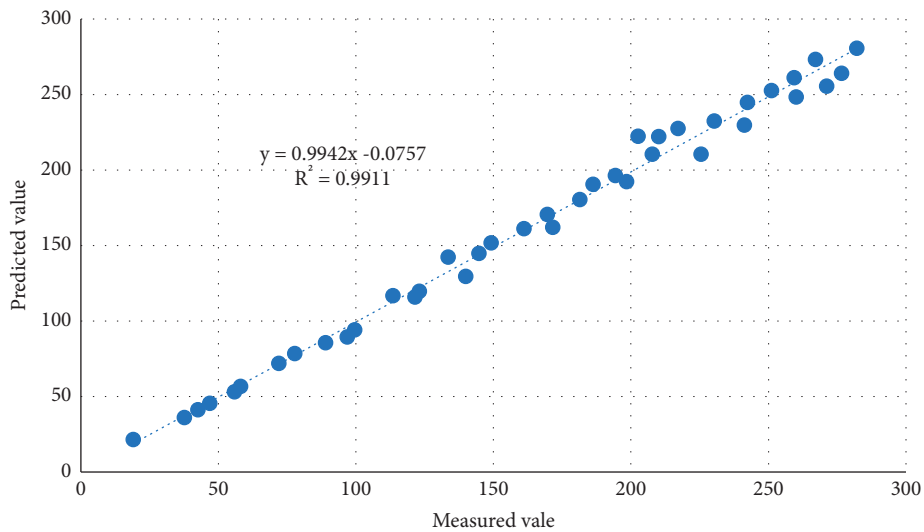


FIGURE 3: Fitting effect of the fresh corn biomass network model.

3.2. *Spatial Distribution of Biomass.* Previous practice and research show that the distribution of fresh corn is significantly related to climate and topographic changes. As a typical species, the distribution of maize shows obvious regularity with the change of terrain, and the spatial distribution of fresh maize biomass is also different. Combined with the BP neural network model of fresh corn biomass, the spatial distribution

of corn biomass in the study area was analyzed from three aspects: altitude, slope, and slope direction.

3.2.1. *Altitude.* It can be seen from Figure 4 that the average fresh corn biomass per unit area shows a gradual upward trend with the increase in altitude. Compared with the area

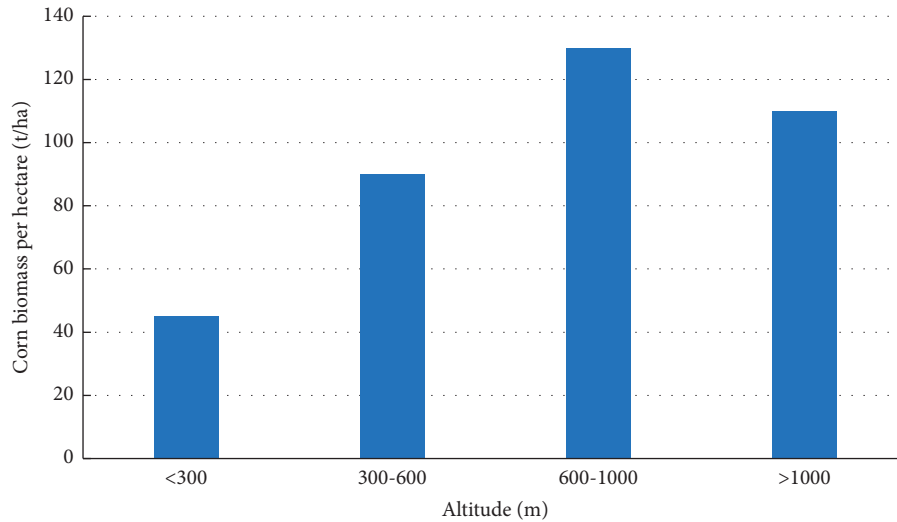


FIGURE 4: Altitude distribution of corn biomass in the study area.

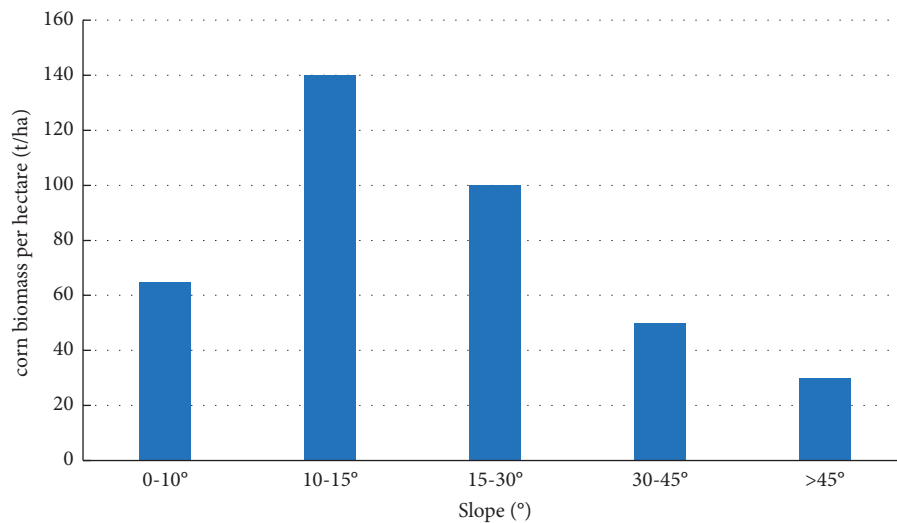


FIGURE 5: Variation of corn biomass level with slope.

below 300 m above sea level, the fresh corn biomass per unit area increased significantly in the area of 300–600 m above sea level, which was related to human production activities. Low-elevation areas tend to be areas with relatively frequent human activities, areas where forestry production is more active, so fresh corn biomass levels are high. Conversely, in areas with relatively high elevations, i.e., 300–1000 meters, the corn quality in this area is higher and the biomass level is lower due to the reduced accessibility of corn fields and relatively less anthropogenic disturbance. In areas above 1,000 meters above sea level, corn biomass levels decreased, possibly due to a decrease in quality due to decreased moisture and climatic conditions due to increased altitude.

3.2.2. Slope. The energy exchange and material circulation of fresh corn ecosystems are obviously affected by the slope. Some scholars have pointed out that the solar radiation

energy, soil bedrock, soil organic matter content, corn water holding capacity, and so on in the ecosystem all show variability within different slope ranges. In addition, increasing the slope can limit human movement. All these reasons lead to different distribution laws of corn biomass on different slopes. The slope is divided into five grades: 0–10° is a flat slope, 10–15° is a gentle slope, 15–30° is a normal slope, 30–45° is a steep slope, and above 45° is a steep slope.

It can be seen from Figure 5 that the average corn biomass per hectare is only 65 t/ha in the flat slope area with a slope of less than 15°. Gentle slopes and slopes are areas with high corn biomass levels, with an average of 140 t/ha for gentle slopes and 100 t/ha and for slope areas and steep slopes, the biomass level decreased sharply. The change trend of biomass level with slope can be attributed to several reasons. First, most of the flat slope areas are at lower altitudes, where human activities are the most frequent and corn is greatly affected by human disturbance and human production activities, so the quality of corn is

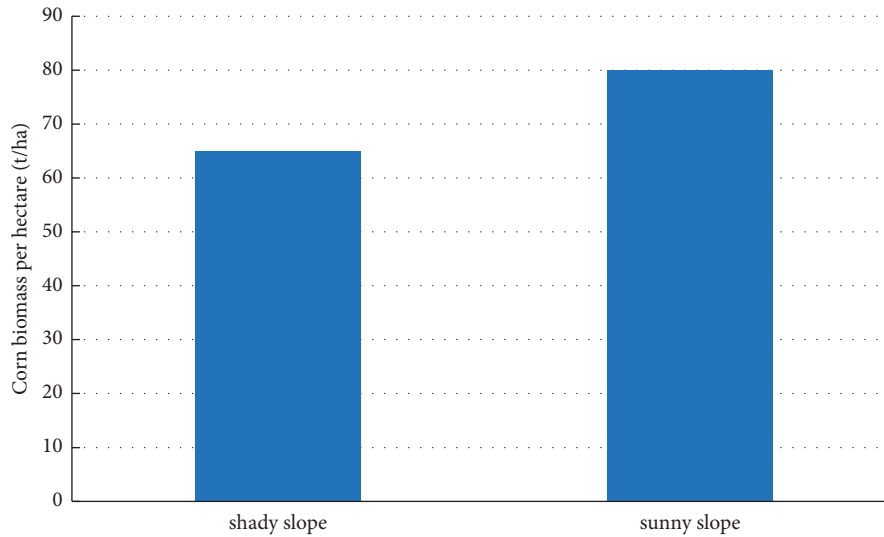


FIGURE 6: Variation of corn biomass level with slope direction.

not high. The biomass level is low. Secondly, gentle slopes and normal slope areas are usually the areas that the government and forest farmers pay more attention to, and they are the objects that people focus on tending and protecting, and as the slope increases, the corn is less damaged by humans, so corn grows. As the slope continued to increase, the climatic conditions decrease, the surface interception decreases, and the ability of corn to maintain soil and water decreases, resulting in the decline of forest quality and the gradual decline of the corn biomass level.

3.2.3. Slope Aspect. Under the same slope and altitude, slope aspect is a factor that has an obvious impact on corn distribution and corn biomass also varies significantly with slope aspect. It is evident from Figure 6 that corn grown on sunny slopes had higher biomass levels than corn grown on shady slopes.

4. Conclusions

- (1) The LM algorithm is used as the training algorithm of the corn biomass BP network model, and a 3-layer BP neural network model is constructed. Among them, the hyperbolic tangent function is selected as the transfer function from the input layer to the hidden layer, the value range is $(-1, 1)$, and the linear function is selected as the transfer function from the hidden layer to the output layer, without any transformation of the output value. By fitting the effect diagram of the BP neural network model, the results show that the predicted value is highly correlated with the measured value, indicating that the fitting effect of the model is ideal.
- (2) We input the image data after principal component transformation and normalization into the BP neural network model, from the three aspects of

altitude, slope, and slope aspect, combined with the BP neural network model of corn biomass, to analyze the distribution of corn biomass in the study area. The results showed that when the altitude was lower than 1000 m, the average biomass per unit area of maize increased with the increase in altitude, and when the altitude was higher than 1000 m, the average biomass per unit area of maize decreased rapidly with the increase in altitude. In terms of slope, in the flat slope area with slope less than 15° , the average biomass of corn per hectare increases rapidly with the increase in slope, and the biomass level of slope is higher than that of steep slope, and the steep slope is unstable. With the increase in slope, the biomass level shows a sharp downward trend. The biomass of maize on sunny slope was higher than that on shady slope.

- (3) The training algorithm of the BP network model is suitable for the construction of the remote sensing model of fresh corn biomass.

Data Availability

The figures and tables used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by Public Service Technology Application Research of Zhejiang Province (GN20C020002) and Zhejiang Science and Technology Major Program on Agricultural New Variety Breeding (2021C02064-4). The authors would like to show sincere thanks to those techniques which have contributed to this research.

References

- [1] L. Han, G. Yang, H. Dai et al., “Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data,” *Plant Methods*, vol. 15, no. 1, pp. 1–19, 2019.
- [2] L. Wang, X. Zhu, Z. Dong, and W. Guo, “Estimation of biomass in wheat using random forest regression algorithm and remote sensing data,” *The Crop Journal*, vol. 4, no. 3, pp. 212–219, 2016.
- [3] X. Jin, Z. Li, H. Feng, Z. Ren, and S. Li, “Deep neural network algorithm for estimating maize biomass based on simulated Sentinel 2A vegetation indices and leaf area index,” *The Crop Journal*, vol. 8, no. 1, pp. 87–97, 2020.
- [4] Y. Uno, S. O. Prasher, R. Lacroix et al., “Artificial neural networks to predict corn yield from compact airborne spectrographic imager data,” *Computers and Electronics in Agriculture*, vol. 4, no. 2, pp. 149–161, 2005.
- [5] W. Zhu, Z. Sun, J. Peng et al., “Estimating maize above-ground biomass using 3D point clouds of multi-source unmanned aerial vehicle data at multi-spatial scales,” *Remote Sensing*, vol. 11, no. 22, p. 2678, 2019.
- [6] S. Khaki, H. Pham, and L. Wang, “Simultaneous corn and soybean yield prediction from remote sensing data using deep transfer learning,” *Scientific Reports*, vol. 11, no. 1, pp. 1–14, 2021.
- [7] Y. Xiaohua, W. Fumin, J. F. Huang et al., “Comparison between radial basis function neural network and regression model for estimation of rice biophysical parameters using remote sensing,” *Pedosphere*, vol. 19, no. 2, pp. 176–188, 2009.
- [8] M. Yoosefzadeh-Najafabadi, D. Tulpan, and M. Eskandari, “Using hybrid artificial intelligence and evolutionary optimization algorithms for estimating soybean yield and fresh biomass using hyperspectral vegetation indices,” *Remote Sensing*, vol. 13, no. 13, p. 2555, 2021.
- [9] S. S. Panda, D. P. Ames, and S. Panigrahi, “Application of vegetation indices for agricultural crop yield prediction using neural network techniques,” *Remote Sensing*, vol. 2, no. 3, pp. 673–696, 2010.
- [10] L. P. Osco, A. P. M. Ramos, E. A. S. Moriya et al., “Modeling hyperspectral response of water-stress induced lettuce plants using artificial neural networks,” *Remote Sensing*, vol. 11, no. 23, p. 2797, 2019.
- [11] P. Sharma, L. Leigh, J. Chang, M. Maimaitijiang, and M. Caffè, “Above-ground biomass estimation in oats using UAV remote sensing and machine learning,” *Sensors*, vol. 22, no. 2, p. 601, 2022.
- [12] M. Shibayama and T. Akiyama, “Seasonal visible, near-infrared and mid-infrared spectra of rice canopies in relation to LAI and above-ground dry phytomass,” *Remote Sensing of Environment*, vol. 27, no. 2, pp. 119–127, 1989.
- [13] M. Yuan, L. L. Huang, J. H. Chen, J. Wu, and Q. Xu, “The emerging treatment landscape of targeted therapy in non-small-cell lung cancer,” *Signal Transduction and Targeted Therapy*, vol. 4, no. 1, pp. 1–14, 2019.
- [14] J. Yue and J. M. López, “Understanding MAPK signaling pathways in apoptosis,” *International Journal of Molecular Sciences*, vol. 21, no. 7, p. 2346, 2020.
- [15] C. L. Dougherty and K. C. Cavanaugh, “Mapping coastal wetland biomass from high resolution unmanned aerial vehicle (UAV) imagery,” *Remote Sensing*, vol. 11, no. 5, p. 540, 2019.
- [16] M. Carranza-García, J. García-Gutiérrez, and J. C. Riquelme, “A framework for evaluating land use and land cover classification using convolutional neural networks,” *Remote Sensing*, vol. 11, no. 3, p. 274, 2019.
- [17] C. P. J. Maury, K. Alii, and M. Baumann, “Finnish hereditary amyloidosis: amino acid sequence homology between the amyloid fibril protein and human plasma gelsoline,” *FEBS Letters*, vol. 260, no. 1, pp. 85–87, 1990.
- [18] C. D. DiNardo, K. Pratz, V. Pullarkat et al., “Venetoclax combined with decitabine or azacitidine in treatment-naïve, elderly patients with acute myeloid leukemia,” *Blood: The Journal of the American Society of Hematology*, vol. 133, no. 1, pp. 7–17, 2019.
- [19] Z. Li, Y. Yi, X. Luo et al., “Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis,” *Journal of Medical Virology*, vol. 92, no. 9, pp. 1518–1524, 2020.
- [20] D. J. Kachamba, H. O. Ørka, T. Gobakken, T. Eid, and W. Mwase, “Biomass estimation using 3D data from unmanned aerial vehicle imagery in a tropical woodland,” *Remote Sensing*, vol. 8, no. 11, p. 968, 2016.
- [21] S. Xia, L. Yan, W. Xu et al., “A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike,” *Science Advances*, vol. 5, no. 4, Article ID eaav4580, 2019.
- [22] M. Marshall and P. Thenkabail, “Developing in situ non-destructive estimates of crop biomass to address issues of scale in remote sensing,” *Remote Sensing*, vol. 7, no. 1, pp. 808–835, 2015.
- [23] U. Kizil, L. Genc, M. Inalpulat, D. Şapolyo, and M. Mirik, “Lettuce (*Lactuca sativa* L.) yield prediction under water stress using artificial neural network (ANN) model and vegetation indices,” *Zemdirbyste Agriculture*, vol. 99, no. 4, pp. 409–418, 2012.
- [24] J. Bendig, A. Bolten, S. Bennertz, J. Broscheit, S. Eichfuss, and G. Bareth, “Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging,” *Remote Sensing*, vol. 6, no. 11, pp. 10395–10412, 2014.
- [25] S. L. James, D. Abate, K. H. Abate et al., “Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017,” *The Lancet*, vol. 392, no. 10159, pp. 1789–1858, 2018.
- [26] B. Liu, Z. Liu, T. Zhang, and T. Yuan, “Non-differentiable saddle points and sub-optimal local minima exist for deep ReLU networks,” *Neural Networks*, vol. 144, pp. 75–89, 2021.
- [27] W. Wu, A. Wang, and M. Liu, “Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China,” *Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.
- [28] K. Khun, N. Tremblay, B. Panneton et al., “Use of oblique RGB imagery and apparent surface area of plants for early estimation of above-ground corn biomass,” *Remote Sensing*, vol. 13, no. 20, p. 4032, 2021.
- [29] P. F. Chen, T. Nicolas, J. H. Wang, V. Philippe, W. J. Huang, and B. G. Li, “New index for crop canopy fresh biomass estimation,” *Spectroscopy and Spectral Analysis*, vol. 30, no. 2, pp. 512–517, 2010.
- [30] C. L. Mohler, “Effects of tillage and mulch on weed biomass and sweet corn yield,” *Weed Technology*, vol. 5, no. 3, pp. 545–552, 1991.
- [31] A. Ngie, F. Ahmed, and K. Abutaleb, “Remote sensing potential for investigation of maize production: review of literature,” *South African Journal of Geology*, vol. 3, no. 2, pp. 163–184, 2014.

- [32] S. Lek and J. F. Guégan, “Artificial neural networks as a tool in ecological modelling, an introduction,” *Ecological Modelling*, vol. 120, no. 2-3, pp. 65–73, 1999.
- [33] G. P. Guo and G. C. Guo, “Quantum secret sharing without entanglement,” *Physics Letters A*, vol. 310, no. 4, pp. 247–251, 2003.
- [34] P. T. Nelson, W. X. Wang, and B. W. Rajeev, “MicroRNAs (miRNAs) in neurodegenerative diseases,” *Brain Pathology*, vol. 18, no. 1, pp. 130–138, 2008.
- [35] J. Wei, H. Cheng, B. Fan, Z. Tan, L. Tao, and L. Ma, “Research and practice of “one opening-one closing” productivity testing technology for deep water high permeability gas wells in South China Sea,” *Fresenius Environmental Bulletin*, vol. 29, no. 10, pp. 9438–9445.
- [36] Q. Qin, H. Cheng, M. Wang, M. Sun, and L. Zhao, “Analyzing the wettability of tight sandstone of taiyuan formation in shenfu block, eastern margin of ordos basin,” *IOP Conference Series: Earth And Environmental Science*, vol. 671, no. 1, Article ID 012022, 2021.

Research Article

The Research of Chinese Martial Arts Cross-Media Communication System Based on Deep Neural Network

Yue Su ¹, Jing Tian,² and Xin Zan³

¹Physical Education Department of Tianjin University of Science and Technology, Tianjin, China

²Physical Education of Tianjin Business Vocational College, Tianjin, China

³Sports Department of Tianjin Ren'ai College, Tianjin, China

Correspondence should be addressed to Yue Su; suyue@tust.edu.cn

Received 29 March 2022; Revised 21 April 2022; Accepted 25 April 2022; Published 28 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yue Su 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 spread of Chinese martial arts is crucial for the world to understand Chinese culture. If only relying on one transmission method, it will lead to the difference of transmission and its lack of certain real time. This will lead to differences in the understanding of Chinese martial arts, which is also not conducive to the spread of Chinese glorious culture. Cross-media communication technology can solve this communication difference problem very well. The deep neural network method was used to fuse relevant features of Chinese martial arts, and it also analyzes the feasibility of neural network technology in cross-media communication. At the same time, this study uses deep neural network to study the timeliness of Chinese martial arts in the process of cross-media communication. The research results show that the convolutional neural network can effectively extract the characteristics of Chinese martial arts and carry out effective dissemination. However, the hybrid convolutional neural network with temporal features has higher accuracy in extracting Chinese martial arts features. This hybrid convolutional neural network is more conducive to the dissemination of Chinese martial arts through cross-media technology, which can ensure its timeliness. The maximum error of deep neural network technology in predicting Chinese martial arts culture is only 2.67%. This part of the error comes from the action characteristics of Chinese martial arts culture, which shows that neural network technology has good feasibility.

1. Introduction

In the history of the development of Chinese culture and civilization, Chinese martial art has a long history and it is a culture that has been passed down. It is not only a representation of sports, but also a representation of Chinese culture [1, 2]. For a country, the national image is important in the communication of the international community. Every country often has a culture as an image characteristic of national exchanges. The characteristics of national cultural image are often a long-term historical culture, which can represent the development experience and historical culture of this country. Chinese martial art is one of the characteristics of national image [3]. It not only represents the historical inheritance of this culture, but also reflects the development history of Chinese culture and the development

experience of China [4]. For the country, national culture is a symbol that inspires national labor and struggle. It is also a belief of the people of the country. Chinese martial arts culture has experienced thousands of years of development history, and it can inspire the hard work and tenacious fighting spirit of the Chinese [5, 6]. As a Chinese citizen, we must respect our national image culture wherever we are. Chinese martial arts culture is also a way for the people of the world to understand China [7, 8]. A country or even a citizen wants to understand the development process of China, and Chinese martial arts culture is a good representative.

Since Chinese martial arts culture is a representative of Chinese cultural symbols, we need to pass on Chinese martial arts culture in various ways, so that people around the world can understand Chinese martial arts culture [9, 10]. In the early stage of development, China generally

disseminated Chinese martial arts culture through newspapers and the Internet [11]. Whether it is a large-scale sports event or a foreign cultural exchange, China often regards Chinese martial arts as a cultural exchange product. For example, during the Beijing Summer Olympics in 2008, China showed the Chinese martial arts culture to the people of the world through a performance. Although this will not increase the number of gold medals in China, it is also an opportunity to show Chinese martial arts to the world. However, there are certain differences in the development of each country [12, 13]. Even some countries are relatively backward in development, which will lead to the lag of cultural transmission. Chinese martial arts culture is relatively special culture. In the process of dissemination, the dissemination of Chinese martial arts culture tends to produce certain differences [14]. This will lead to differences in the world's understanding of Chinese culture. Chinese martial arts culture contains cultural features such as body movements, images, and language [15, 16]. It is difficult to spread the connotation of Chinese martial arts culture to the world just by relying on one communication method. Cross-media technology is a way of disseminating information through the Internet, newspapers, and television [17]. Chinese martial arts culture will improve the timeliness and accuracy of cultural communication through cross-media technology. However, it also has a certain error rate when collecting the characteristics of Chinese martial arts culture, which makes it difficult to ensure the effective dissemination of Chinese martial arts culture.

In order to solve the problem of collecting the characteristics of Chinese martial arts culture and the real-time dissemination of cross-media technology, deep neural network can be considered to be applied in the dissemination process of Chinese martial arts culture [18, 19]. Deep neural network technology is a technology that has been widely used in recent years. The basis of deep neural network is machine learning algorithm and simple neural network technology, which is mainly due to the limitation of computer performance and graphics processing unit (GPU) computing performance. The amount of parameter calculation of deep neural network is huge, which requires GPU technology for computing performance and high-performance computer storage technology. Deep neural network methods can extract deeper data features by increasing the number of network layers. With the rapid development of data volume in each field, deep neural network technology has been favored by people [20]. In the process of cross-media dissemination of Chinese martial arts culture, it not only needs to ensure the effective extraction of the characteristics of Chinese martial arts culture, but also requires the timeliness of dissemination. It is difficult to make sure that Chinese martial arts culture dissemination has better accuracy and timeliness by traditional network or television. The deep neural network method can effectively extract the characteristics and time characteristics of Chinese martial arts culture, which can ensure the accuracy and timeliness of Chinese martial arts culture in the process of cross-media dissemination.

This study uses the deep neural network method to study the accuracy and timeliness of Chinese martial arts culture in the process of cross-media communication. This paper mainly introduces five aspects. The first part mainly explains the development of Chinese martial arts culture and the necessity of its dissemination. It also introduces the defects in the dissemination of Chinese martial arts culture. The second part mainly introduces the research status of Chinese martial arts culture. The third part mainly introduces the system composition and deep neural network method of the cross-media communication process of Chinese martial arts culture. The deep neural network technology used in this study is convolutional neural network (CNN) and hybrid neural network CNN-LSTM technology. The fourth part uses different statistical parameters to introduce the accuracy and timeliness of deep neural network technology in the process of Chinese martial arts culture cross-media dissemination. The fifth part is the summary part of the article.

2. Related Work

Chinese martial arts culture is more important to Chinese cultural image, and it will have different characteristics in the process of dissemination. Many researchers have also conducted related research on Chinese martial arts culture or the characteristics of martial arts culture. Peng and Zhang [21] believed that, with the increasing frequency of international exchanges, cultural exchanges between countries will gradually increase. It uses the comparative analysis method and the variation method to study the characteristics of Chinese martial arts culture. It also analyzes the influence of Chinese martial arts culture on the communication of Chinese martial arts art. Based on this method, he established a reliable model to evaluate the influence of Chinese martial arts culture on the spread. The results of the study indicate that propaganda will affect the communication of Chinese martial arts culture and it will strengthen the international cultural exchange. Ying [22] regarded traditional Chinese martial arts culture as an excellent traditional culture, but with the development of economy it has faded out of people vision. He used numerical analysis and data processing methods to establish a model of influencing factors of Chinese martial arts. He also studied the satisfaction of Chinese martial arts culture and art masses by using goal planning method and correlation analysis method. Through research, he suggested that Chinese martial arts should receive greater publicity, so that the masses can truly understand the connotation of Chinese martial arts. Guo and Fu [23] analyzed the development background of the exposure of martial arts culture in Guangdong. They also analyzed the manifestations of various martial arts cultures and the dissemination of traditional martial arts art and cultures. The results of the study show that the ideology of Guangdong martial arts culture and art originates from traditional Chinese philosophy. The values and symbols of martial arts culture are the important support for the development of Guangdong martial arts culture. Richards [24] researched the art form of Chinese martial arts and the contradiction between the ideological

concept of Chinese martial arts art and Chinese martial arts art with film as the theme. The findings suggest that martial arts movie stars can influence the ideological change between contemporary and traditional culture and they can deepen the understanding of Chinese martial arts culture. Xu [25] believed that the traditional culture of Chinese martial arts represents China's long history, and it can be regarded as a symbol of Chinese culture by foreigners. He also believes that it has an important relationship with Chinese martial arts culture and the spiritual connotation of spreading martial arts culture. This research takes Chinese hosting of the Chinese martial arts Art and Culture Festival as the research object. He uses the method of time series to establish an international art development model for the development of Chinese martial arts culture. The research results show that Chinese martial arts culture has been widely spread around the world, and Chinese martial arts culture and art are at a low level in the world. An and Hong [26] proposed that Maurice Halbwachs collective memory theory is helpful to understand Chinese understanding of martial arts art culture. They used interdisciplinary knowledge to build a memory spectrum of Chinese martial arts culture. Lv [27] mainly conducted related martial arts culture research based on the competition of Chinese martial arts culture. He found that the strategies and actions of Chinese martial arts culture have a corresponding relationship with each other. He used the competitive decision-making algorithm to build the decision-making algorithm of Chinese martial arts. He also studied the decision-making algorithm of Chinese martial arts culture using neural network technology and novel gradient descent method with added momentum and BP neural network method. This study compares the choice of step size as well as the choice of weight size and learning rate. The research results show that the model proposed in this study improves the scale and running time of Chinese martial arts culture decision-making, and it can help professionals to formulate scientific decision-making strategies. From the above literature review, it can be found that researchers have done a lot of research on the development and dissemination of Chinese martial arts culture; it involves traditional mathematical methods and simple neural network methods. It mainly refers to shallow neural network technology, such as BP neural network technology. These neural network techniques can only map the relationship between input and output. These studies do not involve deep neural network methods for Chinese martial arts cultural propagation tasks. The CNN and CNN-LSTM hybrid neural network technology used in this study can extract the characteristics and temporal characteristics of Chinese martial arts culture, which can help relevant personnel to discover the potential characteristics. It is helpful for the dissemination of Chinese martial arts culture.

3. An Application of Deep Neural Network in Chinese Martial Arts

3.1. The Introduction to Deep Neural Network. A deep neural network is a neural network technique that goes deeper

into layers. It can handle the relationship between more complex data, because the amount of data has grown rapidly with the rapid development of economic globalization. The increase in the amount of data can no longer be handled by a simple neural network. BP neural network is the basis of neural network technology; it has the most basic perceptron and backpropagation mechanism. Deep neural network technology also has the basic process of backpropagation mechanism and perceptron. However, deep neural network technology allows deeper networks to perform related prediction operations through weight sharing. The cross-media communication technology of Chinese martial arts culture involves many features, which requires a deeper network level for feature extraction and prediction tasks. Therefore, deep neural network technology is a necessary method to more accurately predict and extract the characteristics and timeliness in the process of Chinese martial arts culture dissemination [28, 29]. Simple neural network technology does not have a weight sharing mechanism, which will increase the amount of parameter computation for the extraction of Chinese martial arts cultural features. There will be many features involved in the process of Chinese martial arts cultural cross-media dissemination. It is difficult for simple neural network technology to effectively extract this potential information.

3.2. The System Design and CNN Algorithm. Cross-media communication can effectively promote the accuracy and timeliness of the dissemination of Chinese martial arts culture, which can ensure the status of Chinese martial arts culture in world culture and the dissemination of Chinese historical information. This research mainly uses the deep neural network method to study the spread of Chinese martial arts culture. Figure 1 shows the system design scheme of Chinese martial arts culture cross-media dissemination. First of all, this research needs to collect the action features, image features, and language features of Chinese martial arts culture, which is to prepare for the deep neural network technology. These data need to go through the data processing stage, and then the processed Chinese martial arts cultural characteristic data is input into the convolutional neural network (CNN) and the long short-term memory neural network (LSTM) [30, 31]. After the relevant features of Chinese martial arts culture are extracted by CNN and LSTM, the data is output after passing through the output layer. These output data are feature data processed by deep neural network technology. These data are the Chinese martial arts culture data that need to be disseminated. Cross-media communication can be carried out in the form of television, newspapers, and the Internet, and they are interdependent [32]. In this study, the characteristics processed by deep neural network technology will be used to spread Chinese martial arts culture through television, newspapers, and the Internet. These characteristics can be propagated in these three media technologies simultaneously, and there is also a mutual connection between them.

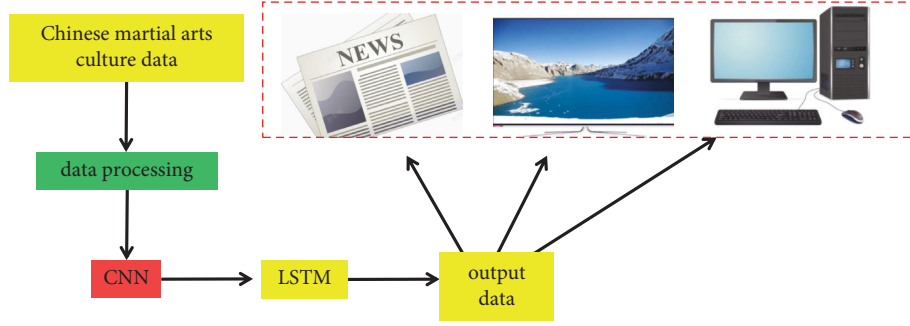


FIGURE 1: The system design of Chinese martial arts cross-media communication.

The deep neural network technology used in this study includes CNN, which is used to extract the spatial characteristics of Chinese martial arts culture, and this technology can also reduce the amount of parameter computation of the deep neural network. Figure 2 shows the computational process of CNN in the extraction of Chinese martial arts cultural features. In Figure 2, the large cube represents the process of feature extraction by the convolutional layer, and each face of the cube represents a feature. The small cube represents the operation process of the pooling layer. These features first need to go through the convolutional layer of CNN to extract the main features; this process will reduce the data volume of features with lower weights. These features are then sampled by pooling layers and nonlinear data processing techniques with activation functions. Finally, the CNN will output the predicted data value through the output layer. The output data will be compared to the actual tag data value. This will involve the backpropagation mechanism of CNN and the error calculation process.

Whether it is a simple BP neural network or a complex deep neural network, it will involve two processes of forward propagation and backpropagation. These two processes involve complex derivation operations. But the derivation operation is solved by automatic differentiation in deep neural networks; equations (1) and (2) show the process of the derivation operation.

$$\Delta\omega_{ji} = -\eta \frac{\partial E}{\partial \omega_{ji}}, \quad (1)$$

$$\Delta u_{ij} = -\eta \frac{\partial E}{\partial u_{ij}}. \quad (2)$$

In the process of backpropagation, CNN will involve many error calculations, which also involve the error calculation of weights and biases. Equations (3) and (4) show a representation of the error calculation.

$$E = \frac{1}{2} \sum_{k=1}^m [d_k - f(\text{net}w_k)]^2 = \frac{1}{2} \sum_{k=1}^m \left[d_k - f\left(\sum_{j=0}^n \omega_{jk} y_j\right) \right]^2, \quad (3)$$

$$E = \frac{1}{2} \sum_{k=1}^m [d_k - f(\text{net}w_k)]^2 = \frac{1}{2} \sum_{k=1}^m \left[d_k - f\left(\sum_{j=0}^n \omega_{jk} y_j\right) \right]^2 = \frac{1}{2} \sum_{k=1}^m \left[d_k - f\left[\left(\sum_{j=0}^n \omega_{jk} f\left(\sum_{i=0}^q u_{ij} \chi_i\right)\right)\right] \right]^2. \quad (4)$$

Error calculations can take many forms, such as mean square error, square root error, etc. Mean squared error is a commonly used method for calculating predicted and actual values. Equation (5) shows the principle of the mean square error calculation.

$$L = \text{MSE}(q^{\text{real}}, q^{\text{pre}}) = \frac{1}{nm} \sum_{k=1}^N \sum_{j=1}^M (q_{kj}^{\text{real}} - q_{kj}^{\text{pre}})^2. \quad (5)$$

The activation function is the key function for nonlinear processing, and equation (6) shows the calculation criterion of the activation function.

$$S(x) = \frac{1}{1 + e^{-x}}. \quad (6)$$

3.3. The Introduction to LSTM Algorithm and Data Processing. The CNN algorithm can only deal with spatial features very effectively, but it is not good at dealing with temporal features. LSTM can memorize some historical information, it has been widely used in the field of speech recognition, and it can handle temporal features very well. The dissemination of the characteristics of Chinese Wushu culture also requires the extraction of temporal

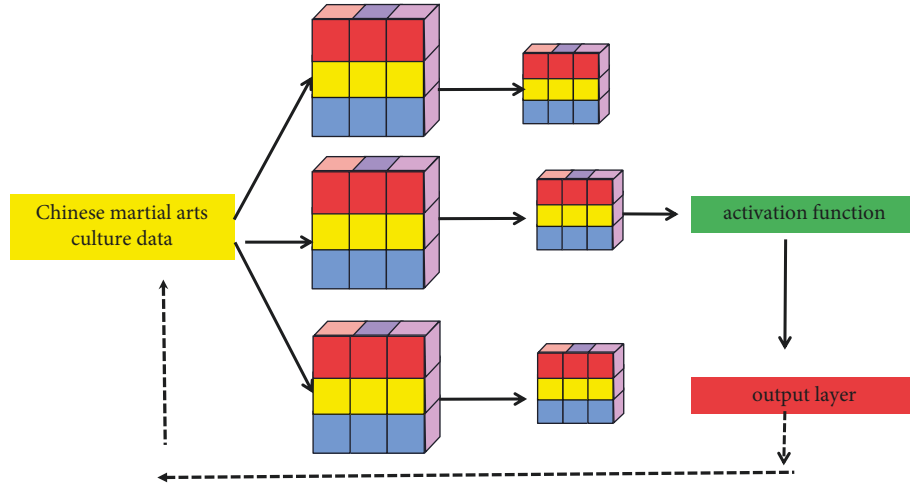


FIGURE 2: The detailed calculation process of CNN.

characteristics, because there are certain differences in the temporal characteristics of Chinese Wushu culture in different periods. Therefore, this study adopts the method of CNN-LSTM hybrid deep neural network to explore the feasibility and accuracy of cross-media communication of Chinese martial arts culture. Figure 3 shows the application process of LSTM in the cross-media communication of Chinese martial arts culture. The input data of LSTM comes from the output data of CNN. This study does not directly provide LSTM with time series data of Chinese martial arts culture. LSTM and CNN have great differences in structure, and its biggest feature is that it has some gate structures. These gate structures can not only input the data, but also input historical state information. It can ensure the preservation of the temporal characteristics of Chinese martial arts culture. The data required by CNN is the data type of input and output, and the data required by LSTM is in the form of time series. The output data of CNN will be input to the input layer of LSTM in the form of time series.

The first layer of LSTM is the input gate structure, which can not only accept the input data of the current state, but also input historical information data. It can give different weights to the current state data and historical information data, which can ensure the integrity of the data and it can retain the time characteristics. Equation (7) shows the calculation criteria for the input gate.

$$f_t = \sigma(w_f \bullet [h_{t-1}, P_t] + b_f). \quad (7)$$

The function of the forget gate is to filter the historical state information, in which a part of the features with a relatively large weight is retained, and a part of the features with a relatively small weight is filtered. This method not only ensures the preservation of the main temporal features, but also ensures that the amount of parameters is not too large. Equation (8) shows how the forget gate is calculated. Equation (9) shows the calculation process of the activation function.

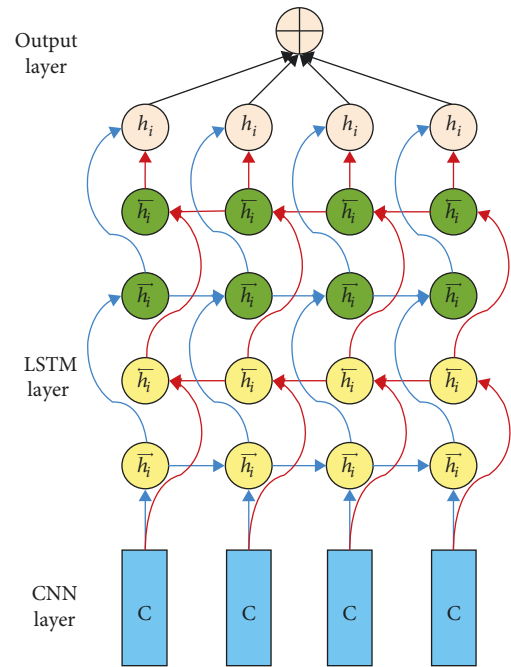


FIGURE 3: The detailed calculation process of LSTM.

$$i_t = \sigma(w_i \bullet [h_{t-1}, P_t] + b_i), \quad (8)$$

$$\bar{C}_t = \tanh(w_c \bullet [h_{t-1}, P_t] + b_c). \quad (9)$$

Equation (10) shows the calculation method of the refresh gate, which assigns different weights to the historical state information and the current state information.

$$\vec{C}_t = f_t \times \vec{C}_{t-1} + i_t \times \bar{C}_t. \quad (10)$$

An output gate is a way to output a time series. Equation (11) shows the calculation of the output gate, which preserves some of the characteristics of the historical state information. Equation (12) shows how the data of the output

gate goes through the nonlinear operation of the activation function.

$$O_t = \sigma\left(w_o \bullet \left[\vec{h}_{t-1}, P_t\right] + b_o\right). \quad (11)$$

$$\vec{h}_t = O_t \times \tanh\left(\vec{C}_t\right). \quad (12)$$

Data processing is a critical step before deep learning training and testing. The characteristics of the cross-media dissemination process of Chinese martial arts culture will involve the movement, image, and language characteristics of Chinese martial arts culture, and these characteristics will be quite different in the collection process. If this study directly uses these data as the input data of CNN, this may cause divergence and lower accuracy in the training process of CNN and LSTM. Therefore, before the training of CNN and LSTM, we need to normalize the characteristic data such as movements and images of Chinese martial arts culture.

4. Result Analysis and Discussion

The data collected in this study come from the relevant data of Chinese martial arts culture performances during the Beijing Olympic Games, which includes the movement characteristics, image characteristics, and language characteristics of Chinese martial arts culture. The cross-media communication process of Chinese martial arts culture will involve the movement, image, and language characteristics of martial arts. The accuracy and timeliness of these characteristics will affect the spread of Chinese martial arts culture. Therefore, this study conducts a related prediction study on three characteristics of Chinese martial arts culture through deep neural network technology. In this study, two methods, CNN and CNN-LSTM, were used to conduct related research. Figure 4 shows the prediction errors of three characteristics of Chinese martial arts culture using the CNN method. In general, the prediction errors of the characteristics of the three Chinese martial arts cultures are all within 3%, which can well guarantee the accuracy of Chinese martial arts culture in the process of cross-media dissemination. The largest error is only 2.78%, and this part of the error comes from the movement characteristics of Chinese martial arts culture. Action characteristics are difficult to be effectively captured by cross-media technology, and it is also difficult to ensure the accuracy of Chinese martial arts cultural characteristics. However, CNN effectively captures the action characteristics of Chinese martial arts culture. The smallest error is only 1.94%, and this part of the error comes from the image features of Chinese martial arts culture.

The parameters of CNN mainly include 5 layers of convolution layers and pooling layers, the number of filters is 128, and the stride is set to 1. The number of layers of CNN-LSTM includes 5 layers of CNN layers and 5 layers of LSTM layers, and the last layer is fully connected layer placed after the LSTM layer. Chinese martial arts culture is a characteristic with time characteristics. In order to verify the superiority of CNN-LSTM, this study compares the error of

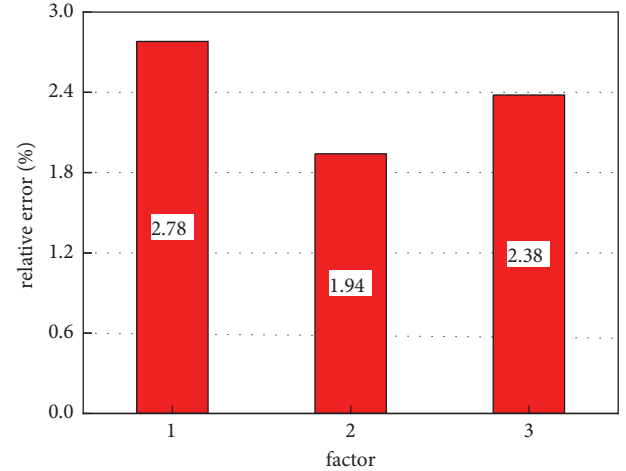


FIGURE 4: The characteristic prediction error of Chinese martial arts culture by CNN method.

CNN method and CNN-LSTM in predicting the cultural characteristics of Chinese martial arts. Figure 5 shows the prediction errors of Chinese martial arts cultural features by the CNN-LSTM method. From Figure 5, it can be intuitively seen that the prediction errors of the three Chinese martial arts cultural characteristics are relatively small compared to the errors of the CNN method, which shows that the CNN-LSTM method is more superior in predicting the Chinese martial arts cultural characteristics. The error distribution of the three features such as action and image is also consistent with the error obtained by the CNN method. The largest error is only 2.67%. This part of the error is also derived from the movement characteristics of Chinese martial arts culture, and this part of the error has been reduced. Language features are data with obvious temporal characteristics, and this part of the error has been reduced to 1.72%.

The largest error in the cultural characteristics of Chinese martial arts comes from the prediction of action characteristics. This study selects the movement characteristics of Wushu alone for analysis. Figure 6 shows the prediction error distribution map of the action features of Chinese martial arts culture using the CNN-LSTM method. It can be seen from Figure 6 that all action errors are distributed within 4%. The prediction error distribution of the action characteristics of Chinese martial arts culture is relatively uniform. There is a large error difference on the left side of the image, but a small error distribution on the right side of the image. This may be due to the superior performance of CNN-LSTM gradually revealed over time. The action features of Chinese martial arts culture are features that have a clear relationship with temporal features, which requires good performance of CNN-LSTM. In general, the action features of Chinese martial arts culture will be better captured by the CNN-LSTM method. If the action features can be captured accurately and in real time, it will be beneficial to the collection and dissemination of cross-media technology.

Image feature is a feature with relatively small prediction error in the process of cross-media communication of Chinese martial arts culture. Figure 7 shows the distribution

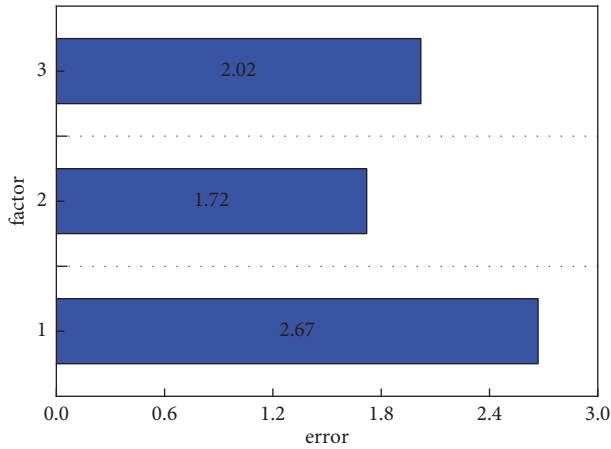


FIGURE 5: The characteristic prediction error of Chinese martial arts culture by CNN-LSTM method.

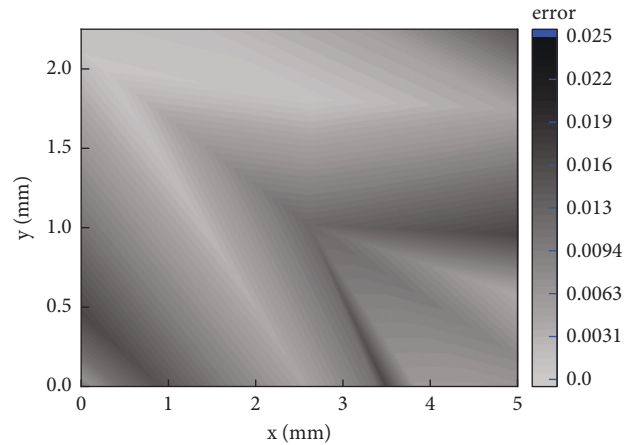


FIGURE 7: The prediction error distribution of image features by CNN-LSTM.

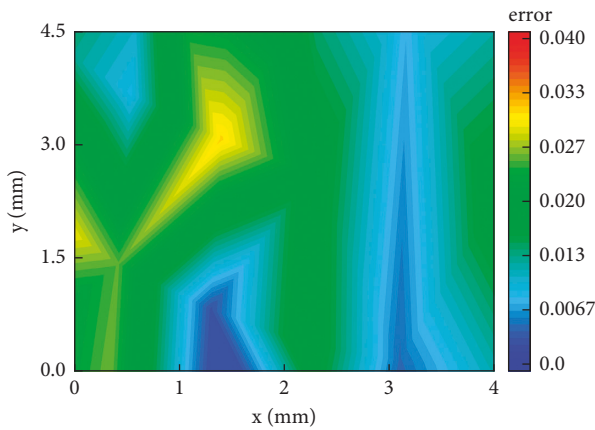


FIGURE 6: The action feature prediction error distribution by CNN-LSTM method.

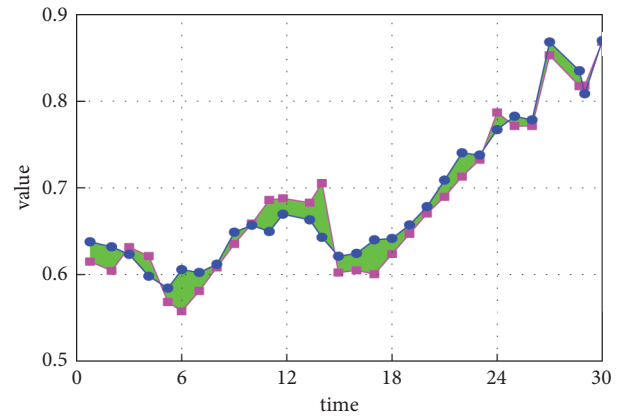


FIGURE 8: The language prediction distribution of Chinese martial arts culture.

of prediction errors for image features of Chinese martial arts culture. The prediction error of image features is different from the distribution of action prediction error, and its distribution is relatively uniform. Moreover, the image prediction errors are basically within 2%. This is more favorable for cross-media communication. The images of Chinese martial arts culture can be spread through newspapers or TV, and this prediction error can well preserve the basic characteristics of Chinese martial arts culture. The distribution of the larger gradient of the image prediction error is mainly in the middle area of the image, which is because the middle area of the image has more characteristics of Chinese martial arts culture. However, in general, image feature prediction of Chinese martial arts culture can ensure effective dissemination through newspapers or television.

The language feature of Chinese martial arts culture is one of the most temporal features. Whether it is spread through TV or the Internet, it needs to ensure the accuracy of the language characteristics of Chinese martial arts culture. Figure 8 shows the distribution curve of language feature prediction of Chinese martial arts culture. Blue

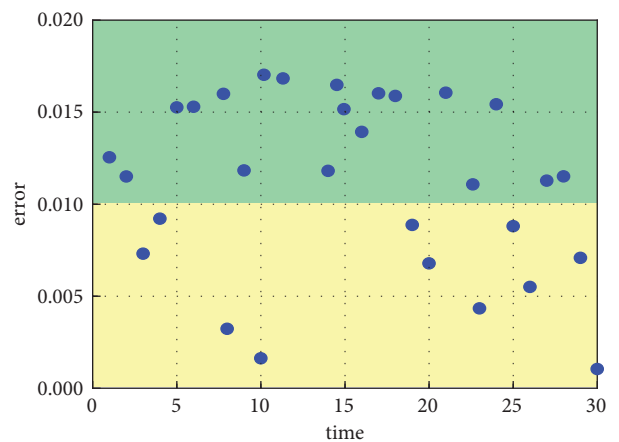


FIGURE 9: The scatter distribution of language prediction errors in Chinese martial arts culture.

represents the actual language feature data, and red represents the predicted language feature data. In general, the predicted values of language features are in good agreement with the actual data values. In the early stage of language

prediction, there is a large error; that is, the area of the green area is relatively large. But over time, the prediction error of the language features of Chinese martial arts culture becomes smaller. Figure 9 shows the scatter plot of the prediction errors of Chinese martial arts culture and language features. The prediction error of all the language features of Chinese martial arts culture is within 2%, which further illustrates the reliability of CNN-LSTM for the cross-media communication of Chinese martial arts culture. Larger language prediction errors still appear in the early stage of Chinese martial arts language features.

5. Conclusions

Chinese martial arts culture not only represents a kind of sports strength, but also reflects China's history and culture as well as China's development history. For the country, Chinese martial arts culture is a national cultural symbol. For citizens, Chinese martial arts culture is an image force that inspires their struggle. The accurate and real-time dissemination of Chinese martial arts culture is an important aspect for both the country and the citizens. There are big differences in the development of science and technology in each country, which will result in differences in the dissemination of Chinese martial arts culture. Cross-media technology is a communication technology that integrates multiple media, which can efficiently spread the movements, images, and language of Chinese martial arts. However, the feature extraction of Chinese martial arts culture is difficult for cross-media technology, and it is difficult to ensure the accuracy and real-time nature of Chinese martial arts cultural features. This study uses deep neural network technology to conduct related research and discussion on this issue.

This study uses two methods, CNN and CNN-LSTM, to conduct prediction research on the relevant characteristics of Chinese martial arts culture in the process of cross-media communication. The research results show that both CNN and CNN-LSTM methods can predict the movement, image, and language features of Chinese martial arts culture well. However, CNN-LSTM has better performance compared to CNN method because it can capture the temporal features of Chinese martial arts culture. The largest prediction error is only 2.67%, and this part of the error comes from the movement characteristics of Chinese martial arts culture. The smallest error is 1.72%, which is a relatively reliable prediction result for the cross-media communication of Chinese martial arts culture. As for the movement prediction error of Chinese martial arts culture, its main error is mainly distributed in the early stage of martial arts culture dissemination. Over time, this part of the error gradually decreases. This further shows that there is a strong time characteristic in the process of cross-media dissemination of Chinese martial arts culture. The distribution of language feature predictions of Chinese martial arts culture is relatively uniform. In general, deep neural network technology has strong accuracy and feasibility in the process of cross-media dissemination of Chinese martial arts culture.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors acknowledge Tianjin Philosophy and Social Science Planning Project, Research on the Theory and Practice of Tai Chi Intercultural Communication (no. TJTYQN20-001).

References

- [1] V. Romanenko, L. Podrigalo, W. J. Cynarski et al., "A comparative analysis of the short-term memory of martial arts athletes of different levels of sportsmanship," *Journal of martial arts anthropology*, vol. 20, no. 3, pp. 18–24, 2020.
- [2] A. H. Gross, R. Feldman, O. Zagoory-Sharon, and Y. Rassovsky, "Hormonal reactivity during martial arts practice among high-risk youths," *Psychoneuroendocrinology*, vol. 121, Article ID 104806, 2020.
- [3] R. B. Santos-Junior, A. C. Utter Jr, S. R. McNulty et al., "Weight loss behaviors in Brazilian mixed martial arts athletes," *Sport Sciences for Health*, vol. 16, no. 1, pp. 117–122, 2020.
- [4] Q. S. Han, M. Theeboom, and D. Zhu, "Chinese martial arts and the olympics: analysing the policy of the international wushu federation," *International Review for the Sociology of Sport*, vol. 11, no. 1, pp. 10–21, 2020.
- [5] C. Hongtu and Z. Cai, "The formal evolution of Chinese martial arts action movies and the spread of core values," *The frontiers of society, science, and technology*, vol. 2, no. 4, 2020.
- [6] T. Aung, S. Goetz, J. Adams et al., "Low fundamental and formant frequencies predict fighting ability among male mixed martial arts fighters," *Scientific Reports*, vol. 11, no. 1, 2021.
- [7] S. J. Stellpflug, W. H. Menton, and R. C. LeFevre, "Analysis of the fight-ending chokes in the history of the Ultimate Fighting Championship mixed martial arts promotion," *The Physician and Sportsmedicine*, vol. 31, no. 1, 2020.
- [8] Y. Tian, "Research on the training path of innovative talents of martial arts and traditional national sports in the new era," *International journal of education and economics*, vol. 3, no. 4, 2020.
- [9] H. Miyata, D. Kobayashi, S. Akifumi, H. Motoike, and S. Akatsuka, "Mindfulness and psychological health in practitioners of Japanese martial arts: a cross-sectional study," *BMC sports science, medicine & rehabilitation*, vol. 12, no. 1, 2020.
- [10] M. Theeboom, D. Zhu, and J. Vertonghen, "'Wushu belongs to the world.' But the gold goes to China. . . : the international development of the Chinese martial arts," *International Review for the Sociology of Sport*, vol. 52, no. 1, pp. 3–23, 2017.
- [11] P. Pawelec, P. Świder, and W. J. Cynarski, "Martial arts tourism: meta-analysis of publications from 2005-2020," *Sustainability*, vol. 12, no. 17, p. 7069, 2020.
- [12] A. A. Figueiredo, W. Błach, Z. Bujak, R. J. Maroteaux, and W. J. Cynarski, "Martial arts tourism of the 'Europe-Far east' direction, in the opinion of grand masters," *Sustainability*, vol. 12, no. 21, p. 8857, 2020.

- [13] Y. Jiang, D. Hu, and J. Hu, "Based on goal programming method of traditional Chinese martial arts culture protection and inheritance," *Journal of Computational and Theoretical Nanoscience*, vol. 13, no. 12, Article ID 10025, 2016.
- [14] J. Zhang, "A cultural discourse analysis to Chinese martial arts movie in the context of glocalization: taking crouching tiger, hidden dragon and hero as cases," *Advances in Language and Literary Studies*, vol. 10, no. 3, pp. 32–41, 2019.
- [15] Q. Yajun, "Evolution and prospect of Chinese martial arts during the rising course of China," in *Proceedings of the 2016 2nd International Conference on Economics, management engineering and education technology*, pp. 1121–1123, Sanya, China, November 2016.
- [16] S. S. M. Fong, J. S. M. Chan, Y. H. Bae et al., "Musculoskeletal profile of middle-aged Ving Tsun Chinese martial art practitioners: a cross-sectional study," *Medicine*, vol. 96, p. 5961, 2017.
- [17] I. Ma, "Between historicity and fictionality: Xiang kairan, martial arts fiction, and Chinese narrative tradition," *Modern Chinese Literature and Culture*, vol. 31, no. 1, pp. 229–263, 2019.
- [18] Z. Ai, "Quantitative CT study of martial arts sports injuries based on image quality," *Journal of Visual Communication and Image Representation*, vol. 60, no. 4, pp. 417–425, 2019.
- [19] J. Yang, "Martial arts cinema in civil society of postcolonial Hong Kong: kung fu hustle and the grandmaster," *Critical Arts*, vol. 34, no. 6, pp. 28–40, 2020.
- [20] Y. Jia, M. Theeboom, and D. Zhu, "Teaching traditional Chinese martial arts to con-temporary Chinese youth—a qualitative study with youth wushu coaches in China," *Archives of budo*, vol. 16, no. 7, pp. 1–10, 2020.
- [21] X. L. Peng and Q. W. Zhang, "Influence factors and improvement suggestions of Chinese wushu communication in Chinese martial arts," *Ekoloji*, vol. 28, no. 107, pp. 2755–2760, 2019.
- [22] L. Ying, "New ecology: Chinese traditional martial arts culture protection and carry forward game analysis," *Ekoloji*, vol. 28, no. 107, pp. 2773–2778, 2019.
- [23] Y. Gao and C. K. Fu, "The culture of Guangdong martial arts," *Sage Open*, vol. 9, no. 3, Article ID 2158244019861459, 2019.
- [24] S. Richards, "The patriotic narrative of Donnie Yen: how martial arts film stars reconcile Chinese tradition and modernity," *Celebrity Studies*, vol. 10, no. 2, pp. 276–284, 2019.
- [25] F. Xu, "Dissemination and inheritance of wushu spirit in the context of cultural globalization and its social environment," *Ekoloji*, vol. 28, no. 107, pp. 2867–2772, 2019.
- [26] L. An and F. Hong, "Body center dot experience center dot imagination: the collective memory of Chinese martial arts," *International Journal of the History of Sport*, vol. 35, no. 15, pp. 1588–1602, 2018.
- [27] H. Lv, "Martial arts competitive decision-making algorithm based on improved BP neural network," *Journal of healthcare engineering*, vol. 7, no. 10, Article ID 9920186, 2021.
- [28] H. Fu, J. Cheng, Y. Xu et al., "Disc-aware ensemble network for glaucoma screening from fundus image," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 2493–2501, 2018.
- [29] K. A. Mills, "What are the threats and potentials of big data for qualitative research?" *Qualitative Research*, vol. 18, no. 6, pp. 591–603, 2018.
- [30] A. Mohamed, M. K. Najafabadi, Y. B. Wah, E. A. K. Zaman, and R. Maskat, "The state of the art and taxonomy of big data analytics: view from new big data framework," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 989–1037, 2020.
- [31] Y. Wang and F. Yang, "A fractional-order CNN hyperchaotic system for image encryption algorithm," *Physica Scripta*, vol. 96, no. 3, Article ID 35209, 2021.
- [32] A. Zhang and N. Lv, "Research on the impact of big data capabilities on government's smart service performance: empirical evidence from China," *IEEE Access*, vol. 9, no. 4, Article ID 50523, 2021.

Retraction

Retracted: CT-ML: Diagnosis of Breast Cancer Based on Ultrasound Images and Time-Dependent Feature Extraction Methods Using Contourlet Transformation and Machine Learning

Computational Intelligence and Neuroscience

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

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

References

- [1] B. Hajipour Khire Masjidi, S. Bahmani, F. Sharifi, M. Peivandi, M. Khosravani, and A. Hussein Mohammed, "CT-ML: Diagnosis of Breast Cancer Based on Ultrasound Images and Time-Dependent Feature Extraction Methods Using Contourlet Transformation and Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1493847, 15 pages, 2022.

Research Article

CT-ML: Diagnosis of Breast Cancer Based on Ultrasound Images and Time-Dependent Feature Extraction Methods Using Contourlet Transformation and Machine Learning

Behnam Hajipour Khire Masjidi,¹ Soufia Bahmani ,^{2,3} Fatemeh Sharifi,³ Mohammad Peivandi,⁴ Mohammad Khosravani,⁵ and Adil Hussein Mohammed⁶

¹Department of Computer Engineering, Islamic Azad University Tehran North Branch, Tehran, Iran

²Department of Computer Engineering and Information Technology, Amirkabir University of Technology, Tehran 15875-4413, Iran

³Department of Electrical Engineering, University of Applied Science and Technology, Bushehr, Iran

⁴Hochschule für Technik und Wirtschaft Berlin (HTW Berlin), Berlin, Germany

⁵Department of Electrical & Computer Engineering, Arak University of Technology, Arak, Iran

⁶Department of Communication and Computer Engineering, Faculty of Engineering, Cihan University-Erbil, Kurdistan Region, Iraq

Correspondence should be addressed to Soufia Bahmani; soufia.bahmani@aut.ac.ir

Received 4 December 2021; Revised 2 February 2022; Accepted 28 February 2022; Published 24 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Behnam Hajipour Khire Masjidi 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.

Breast diseases are a group of diseases that appear in different forms. An entire group of these diseases is breast cancer. This disease is one of the most important and common diseases in women. A machine learning system has been trained to identify specific patterns using an algorithm in a machine learning system to diagnose breast cancer. Therefore, designing a feature extraction method is essential to decrease the computation time. In this article, a two-dimensional contourlet is utilized as the input image based on the Breast Cancer Ultrasound Dataset. The sub-banded contourlet coefficients are modeled using the time-dependent model. The features of the time-dependent model are considered the leading property vector. The extracted features are applied separately to determine breast cancer classes based on classification methods. The classification is performed for the diagnosis of tumor types. We used the time-dependent approach to feature contourlet sub-bands from three groups of benign, malignant, and health control test samples. The final feature of 1200 ultrasound images used in three categories is trained based on k -nearest neighbor, support vector machine, decision tree, random forest, and linear discrimination analysis approaches, and the results are recorded. The decision tree results show that the method's sensitivity is 87.8%, 92.0%, and 87.0% for normal, benign, and malignant, respectively. The presented feature extraction method is compatible with the decision tree approach for this problem. Based on the results, the decision tree architecture with the highest accuracy is the more accurate and compatible method for diagnosing breast cancer using ultrasound images.

1. Introduction

Breast cancer is becoming one of the most severe diseases that affect people worldwide [1]. This condition primarily affects women, although it can also impact men. It is healed by recognizing illnesses early on and curing them. Cancer-related deaths are also on the rise in this area [2].

Consequently, early detection of breast anomalies can lower the mortality rate [3]. In traditional deep learning algorithms, the complicated environment of the feature extraction stage impairs the state's precision and effectiveness [4]. A clinical examination usually carried out by a physician effectively detects a wide range of breast cancer kinds. The doctors first cut segment biopsy samples and then analyze

them with hematoxylin and eosin staining in the first phase of this procedure. Eosin attaches to proteins and emphasizes other components, whereas hematoxylin binds to DNA and accentuates nuclei [5]. Moreover, pathologists examine tissue samples using microscopes to visualize highlighted locations in digital pictures. The assessment of tissue biopsies permits early clues of tissue biopsies to be identified. Experienced pathologists, on the other hand, devote a significant amount of time and effort to this endeavor. A breast cancer diagnosis is a time-consuming and costly procedure. It is highly dependent on the pathologist's past knowledge and the accuracy of histopathology [6]. Examination, mammogram, ultrasonography, magnetic resonance imaging, and positron emission tomography/computed tomography have all been studied for their diagnostic value. There are presently no established diagnostic parameters or reference methodologies for assessing efficacy. Furthermore, a few exploratory investigations utilizing enhanced contrast magnetic resonance imaging, dispersion magnetic resonance, or positron emission high-resolution computed tomography have shown encouraging findings, indicating the need for more study. The ultrasound would have the edge over other scanning technologies in anticipating early tumor responses and developing chemo-switch tactics since it is noninvasive and widely available [7]. Previous research has focused on the association between diagnostic characteristics and molecular subtypes and distinguishing benign and malignant breast cancers using ultrasound pictures. Breast tumors of the triple-negative type had a higher chance of having constricted margins. They had a lower risk of calcifications [7]. The necessity for an adjuvant screening tool has been recognized because of the reduced sensitivity of screening mammography in thick breasts, and ultrasonography has been proposed as a viable supplementary screening modality [7]. Even though ultrasonography is frequently used as a supportive screening technique in Asia [8, 9], there has been little research on the survival advantages of screening ultrasound for breast cancer.

A two-dimensional contourlet is used as the input picture in this work. The time-dependent model is used to represent the sub-banded contourlet coefficients. The primary property vector is made up of the characteristics of the time-dependent model. The collected characteristics are used independently to define breast cancer classifications based on classification algorithms. The categorization is used to determine the kind of tumor. We employed the time-dependent method to highlight contourlet sub-bands from three sets of test samples: benign, malignant, and health control. The outcomes of three modes of classic classification methods, including k-nearest neighbor(KNN), support vector machine(SVM), decision tree(DT), random forest (RF), and linear discrimination analysis (LDA) approaches, are documented, as well as the final feature employed in each.

2. Literature Review

Automated breast cancer diagnosis in mammogram images using moth flame optimization based on the extreme learning machine approach was described by Muduli et al.

[10]. The breast cancer pictures used in this study were acquired from the MIAS collection. The image was then preprocessed to eliminate any noise. The lifting wavelet decomposition was then used to retrieve the features. An extreme learning machine classification was used to classify the images. The moth flame optimization technique optimized the extreme learning machine variables. Breast cancers were categorized as normal or abnormal, benign, or malignant, with 94.76 percent (regular vs. dysfunction) and 97.80 percent (benign vs. malignant) accuracy (benign vs. malignant). The number of features available with this technique is limited. Melekoodappattu and Subbian [11] used a hybrid extreme learning network with the fruit fly optimization classifier algorithm to diagnose breast cancer automatically. Mammography breast cancer images were obtained here to identify breast cancer. The images were then preprocessed to eliminate any noise. The gray level co-occurrence matrix approach was used to feature extraction. The retrieved traits were then used to categorize the pictures as normal, benign, or malignant. The extreme learning machine algorithm was used to classify the images. The fruit fly optimal solution was used to optimize the weight parameters of extreme learning machines. The accuracy of the experimental results is 97.5 percent. The error rate rose as the number of characteristics retrieved grew due to the method's shortcoming.

Sasikala et al. [12] developed a hybrid technique based on the binary firefly method with optimum-path forest classification for detecting breast cancer by merging cranio-caudal and mediolateral oblique views. The GLOBOCAN database was used to provide the initial breast cancer pictures. The images were then preprocessed to eliminate any noise. A local binary pattern was used to extract the picture characteristics. Mediolateral oblique and cranio-caudal aspect mammography were among the characteristics retrieved. Using a hybrid technique based on binary algorithms and an optimum-path forest classifier, these characteristics were combined. The reliability of the approach described is 98.56 percent. Because of the feature fusion procedure, the failure rate rose. Integrating an optimum wavelet statistics structure with the recurrent neural network for tumor identification and tracking was presented by Begum and Lakshmi [13]. The MIAS dataset provided the input mammography images. To eliminate the sounds, the images were preprocessed. The textural characteristics were then extracted. A recurrent neural network classifier was also used to classify the retrieved characteristics. The opposing gravitational search technique improved the existing neural network parameters. The aberrant image was identified and then segmented. The region of interest area was separated using a modified region expanding method. The accuracy of this strategy is 96.43 percent. The given approach includes a false alarm restriction. Fei et al. [14] introduced a doubly supervised factor transference classification to handle transfer learning between unbalanced modalities using labeled data as directed. The suggested method has two algorithms: paired bimodal ultrasound images with shared tags and unpaired pictures with separate labels. Those above used the gradient descent in the support vector machine

plus's specifically designed transfer learning paradigm. In contrast, the latter used the Hilbert–Schmidt autonomy set of criteria for transferring knowledge between the unpaired image data, consisting of single-modal BUS images and EUS images from paired bimodal data. As a result, parameter transfer was used to construct doubly supervised knowledge transfer in a unified optimization problem. The suggested method for the ultrasound-based detection of breast malignancies was tested in two tests. The proposed approach outperformed all comparable algorithms in the experiments, indicating a broad spectrum of uses. Yan et al. [15] developed a peptide MG that targets the tumor-driving protein, MDMX, and causes its destruction. Xu et al. [16] proposed an MTL method for segmenting and categorizing images of the tongue. Our combined strategy is more accurate than the current tongue characterization methods, as demonstrated in the experimental results. A novel feature selection algorithm was used by Tang et al. [17] to identify tissue-specific DNAm at CpG sites. Using a random forest algorithm, we constructed classifiers capable of identifying the origin of tumors with high specificity based on the DNAm profiles of the malignancies.

Zeebaree et al. [18] proposed a features-based fusing approach based on uniform-local binary pattern improvement and filtering noise removal. To overcome the restrictions above and fulfill the study's goal, a new classifier was presented that enriches the local binary pattern characteristics depending on the new threshold. This article introduced a two-stage multilevel fusion technique for the auto-classification of stationary ultrasounds of breast cancer. Using the preprocessing procedure, many pictures were first created from a single image. The median and Wiener filters were used to reduce speckle noise and improve ultrasound visual smoothness. By minimizing the overlap between the benign and malignant picture classes. Second, the fusion technique enabled the creation of various characteristics from diverse filtered pictures. The viability of categorizing ultrasound pictures using the LBP-based structuring element was proven. The suggested approach produced high accuracy (98%), recall (98%), and specificity (98%). Consequently, the fusion procedure, which may assist in generating a robust judgment based on distinct characteristics obtained from different filtered pictures, enhanced the accuracy, sensitivity, and specificity of the new classifier of LBP features. The study by Briganti et al. [19] examined the network structure of alexithymia components and compared the results with relevant prior studies. Rezaei et al. [20] focused on the use of remote sensing methods to generate a geological map of the Sangan area using ASTER satellite imagery. Zhang et al. [21] suggested a privacy-preserving optimization of the clinical pathway query scheme (PPO-CPQ) in order to attain a safe clinical pathway inquiry in e-healthcare. Liu et al. [22] propose a novel perceptual consistency ultrasound image super-resolution (SR) method, which takes only the linear-resolved ultrasound data and guarantees that the generated SR image is consistent with the original LR image, and vice versa. Eslami et al. [23] developed a multi-scale attention-based convolutional neural network for multi-class categorization of road pictures.

Sadeghipour et al. [24] developed a hybrid approach using both a firefly algorithm and an intelligent system to detect breast cancer. Rezaei et al. [25] proposed a data-driven approach to segmenting hand parts on depth maps without the need for extra labeling. Ahmadi et al. [26] proposed a classifier used for diagnosing brain tumors. Based on the results of the ROC curve, the given layer may segregate the brain tumor with a high true-positive rate. Zhang et al. [27] assembled train, test, and exterior test sets using breast ultrasound pictures from two clinics. The training data were used to create an optimal deep learning model. Both the test set and the exterior test set were used to test the validity. Medical experts used the BI-RADS classification to evaluate the clinical outcomes. They classified breast cancer into molecular subgroups based on the expression of the hormone receptor and the female epidermal growth factor receptor.

The deep learning model's capability to identify molecular subtypes was verified in the testing set. In one investigation, the deep learning model was highly influential in detecting breast cancers from ultrasound pictures. As a result, the deep learning model can drastically minimize the number of needless biopsies, particularly in individuals with BI-RADS 4A. Furthermore, this model's prediction capacity for molecular subtypes was good, with therapeutic implications. Table 1 shows the summary of research relate to breast cancer diagnosis and feature extraction methods. Based on the literature review, it can be concluded that some of the feature extraction methods are based on direct analysis of the mammographic or ultrasound image. Moreover, the number of feature extraction method is limited. Therefore, because of the complexity of analyzing ultrasound images, proving a novel method is challenging.

3. Methods and Material

3.1. Contourlet Transformation (CT). It is critical in machine learning to show a picture to extract vital and desirable properties such as the outer boundary. Contourlet transformation is a comparatively recent transformation created to enhance wavelet picture representation. A contourlet filter may be used with many angles at different resolutions, unlike the discrete wavelet processing, which employs just three vertical, horizontal, and diameter filters to extract the appropriate picture components. As a result, the borders of things are retrieved at various angles, referred to as contours. This transformation can give more precise borders than previous edge editing techniques and capabilities like displaying borders at various angles, densities, and ultimate tensile[8]. This transition has two primary steps, as indicated in Figure 1. The Laplacian pyramid is used to scale and find edges and interruptions in the first stage. The directional filter bank is used in the second stage to connect inconsistent locations and form linear structures. The Laplacian pyramid is given to the picture of the low-pass filter first and then eliminated from the main image, leaving the differential image with details and high-frequency elements. After that, factor (2, 2) is used to sample the downstream substrate, and the process is repeated numerous times. The high-frequency

TABLE 1: Summary of research work about breast cancer diagnosis using machine learning methods.

Author	Year	Dataset	Image type	Feature extraction	Classification	Accuracy
Patil & Biradar [28]	2021	MIAS	Mammography	Gray level co-occurrence matrix	Convolutional neural network, recurrent neural network	98.3%
Masud et al. [29]	2021	Rodrigues	Ultrasound	Image	Convolutional neural networks	—
Chung et al. [30]	2021	EHR	Ultrasound	—	Statistical method	—
Fei et al. [14]	2021	Nanjing drum tower hospital	B-mode and elastography ultrasound	Gray level co-occurrence matrix	Doubly supervised parameter transfer classifier	86.73%
Muduli et al. [10]	2020	MIAS	Mammography	Lifting wavelet transform	Extreme learning machine	94.76%
Melekoodappattu et al. [11]	2020	MIAS	Mammography	Gray level co-occurrence matrix	Fruit fly optimization algorithm and extreme learning machine	97.5%
Sasikala et al. [12]	2020	DDSM and INbreast	Mammography	Local binary pattern	Binary firefly approach with optimum-path forest classifier	98.56%
Begum et al. [13]	2020	MIAS	Mammography	Image	Optimal wavelet statistical texture and recurrent neural network	96.43%
Khandezamin et al. [31]	2020	WBCD, WDBC, WPBC	Digitized image of a fine needle aspirate	Logistic regression	Group method data handling neural network	99.4%
Vo et al. [32]	2019	Bioimaging 2015, BreakHis	Histopathology	Image	Incremental boosting convolution networks	96.45%

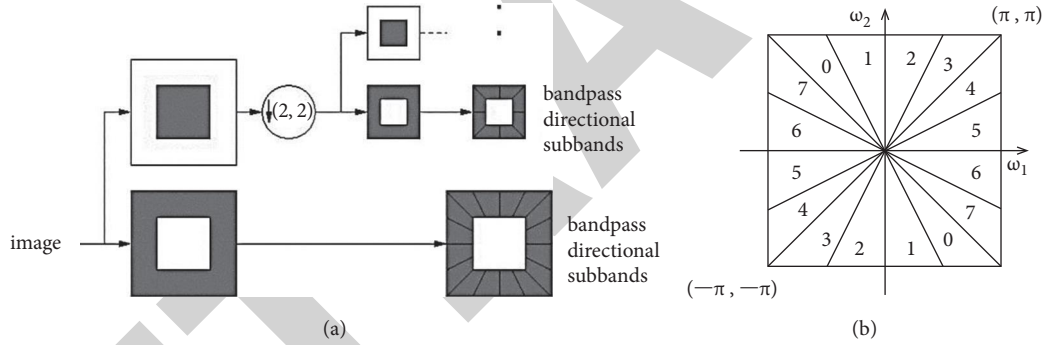


FIGURE 1: The framework of contourlet transformation.

differential image of the directional bank filter is used at each analysis stage to correlate the values on a single scale and separate the overhead sub-band.

This modification yields a set of low-frequency variables that include the highest-level estimation components (lower spatial resolution) as well as a set of high-frequency coefficients that include sensitive components and sharp edges at different scales [33].

3.2. Time-Dependent Feature Extraction Methods. The discrete Fourier transform is supposed to explain the signals trace as a function of frequency $X[k]$ as a product of the sampled depiction of the signal as $x[j]$ with $j = 1, 2, \dots, N$, length N , and sampling rate f_s Hz. Assume we use Parseval's theorem, which states that the whole square of a function is the entire square of its transformation. In that instance, we begin the feature extraction procedure [33]:

$$\sum_{j=0}^{N-1} |x[j]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]X^*[k]| = \sum_{k=0}^{N-1} P[k]. \quad (1)$$

Let $P[k]$ be the phase-excluded power spectrum, per the preceding formula. It implies that multiplying $X[k]$ by the $X^*[k]$ conjugate divided by N yields the frequency index. The whole definition of frequency given by the Fourier transform is commonly understood to be symmetrical concerning zero frequency; that is, it contains equal portions extending to both positive and negative frequencies. This symmetry does not exist throughout the spectrum, including positive and negative frequencies. We cannot use spectral power from the time domain since we have complete access. All irregular moments are also zero by the statistical method of the frequency distribution, according to the idea of a one-minute m of the order n $P[k]$ of the power spectral density [33].

TABLE 2: The equations of root squared moment.

Root squared zero-order moment	Root squared second-order moments	Root squared fourth-order moments
$\overline{m}_0 = \sqrt{\sum_{j=0}^{N-1} x[j]^2} \text{ s}$ $m_0 = \overline{m}_0^\lambda / \lambda$	$\overline{m}_2 = \sqrt{\sum_{k=0}^{N-1} k^2 P[k]} = \sqrt{\sum_{j=0}^{N-1} (\Delta x[j])^2}$ $m_2 = \overline{m}_2^\lambda / \lambda$	$\overline{m}_4 = \sqrt{\sum_{k=0}^{N-1} k^4 P[k]} = \sqrt{\sum_{j=0}^{N-1} (\Delta^2 x[j])^2}$ $m_4 = \overline{m}_4^\lambda / \lambda$

TABLE 3: Time-dependent features of the signal.

Logarithmic features		
$f_1 = \log(m_0)$	$f_2 = \log(m_0 - m_2)$	$f_3 = \log(m_0 - m_4)$
Sparseness		
$f_4 = \log(m_0 / \sqrt{m_0 - m_2} \sqrt{m_0 - m_4})$	Irregularity factor (IF)	
	$f_5 = \sqrt{m_2/m_0} / \sqrt{m_4/m_2}$	Covariance (COV)
		$f_6 = \log(1/\bar{x} \sqrt{\sum_{j=0}^{N-1} (x - \bar{x})^2/n})$
Teager energy operator (TEO):		
	$f_7 = \log(\sum_{j=0}^{N-1} x^2[j] - x[j-1]x[j+1])$	

$$m_n = \sum_{k=0}^{N-1} k^n P[k]. \quad (2)$$

Let $n = 0$ be employed, the Parseval theorem may be applied, and the Fourier transform time-differentiation feature can be used for nonzero amounts of n . For various time signals, such a characteristic explicitly shows that the n' -th equals multiplying the k by the spectrum to the n' -th power, the derivative of a time-domain function referred to as Δ^n :

$$F[\Delta^n x[j]] = k^n X[k]. \quad (3)$$

The root squared zero-order moment is a function that depicts the frequency domain's total power. By separating all channels into zero-order moments, all channels can standardize their corresponding zero-order moments. Also, the root squared second and fourth-order moments are utilized as power, but the frequency functions are referred to by a spectrum shifted $k^2 P[k]$. Because including the second and fourth signals reduces the overall energy of the signal, we use a power transformation to normalize the domain of m_0 , m_2 , and m_4 to reduce the noise effect on all moment-based characteristics. The experimental value of λ is set to 0. As a result of these settings, the top three characteristics retrieved are listed in Table 2 [33].

Table 3 shows the signal's time-dependent characteristics. Sparseness is a metric that estimates the amount of vector energy in only a few additional components based on these equations. Due to differentiation and $\log m_0/m_0 = 0$, a feature indicates a vector with all elements comparable to a zero-sparseness index, that is, m_2 , and, $m_4 = 0$, when it should need a value larger than zero for all other sparseness levels. The irregularity factor expresses the ratio of peak numbers divided by zero crossings up. Their spectral examples can only define the amount of upward zero crossings and the number of peaks in a random signal [33]. The Teager energy operator depicts the size of the signal amplitude and instantaneous changes that are exceptionally responsive to slight changes, and covariance is the ratio of the standard deviation on arithmetic averages. Teager energy operator was first introduced for nonlinear voice signal modeling, but it was later used for signal processing.

3.3. Proposed Feature Extraction Methods. This article wants to employ machine learning algorithms to identify breast cancer. First, we employed contourlet transformation to decompose input images into contourlet sub-bands in this manner. The contourlet pictures obtained are utilized to derive classification features. Then, with the help of nine sub-bands, the time-dependent model is employed to extract features. The principal component analysis (PCA) approach reduces the number of features. Then the extracted feature is utilized to classify breast cancer using multiple machine learning algorithms. Figure 2 shows the block diagram of the proposed method. The following is the pseudocode for the provided method:

3.4. Machine Learning Classification Methods. Machine learning studies automated systems that learn via reasoning and patterning without being explicitly programmed using algorithms models [34]. Over time, machine learning algorithms learn and develop on their own. A support vector machine is a supervised machine learning model for two-group classification issues that employs classification techniques. Support vector machine is a rapid and trustworthy classification technique that works well with small data [35]. SVMs are a collection of supervised learning algorithms for classification and regression issues. A decision tree method divides data into subgroups in a machine learning model. The goal of a decision tree is to condense the training data into the most miniature feasible tree. The decision tree is a supervised linear classifier that performs a split test in its core node and anticipates a target class example in its leaf node [36]. KNN is a feature similarity-based, nonparametric, slow learning method. It is a pattern recognition algorithm that works well. It is a straightforward classifier that categorizes datasets based on the category of their nearest neighbors. KNN is likely to be an excellent choice for a classifying investigation that involves large databases. Healthcare databases include many data; hence, KNN can successfully predict a new sample point class. According to studies, the new dimensionality-decreased KNN classification method surpasses the previous probabilistic neural network scheme in terms of average accuracy, sensitivity,

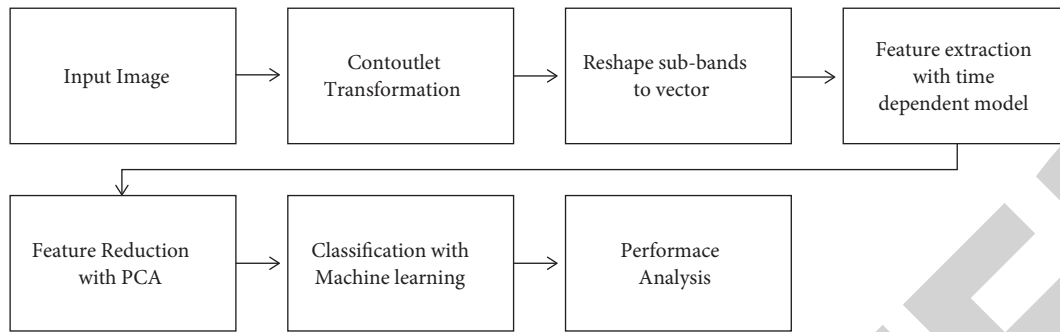


FIGURE 2: The block diagram of the proposed method.

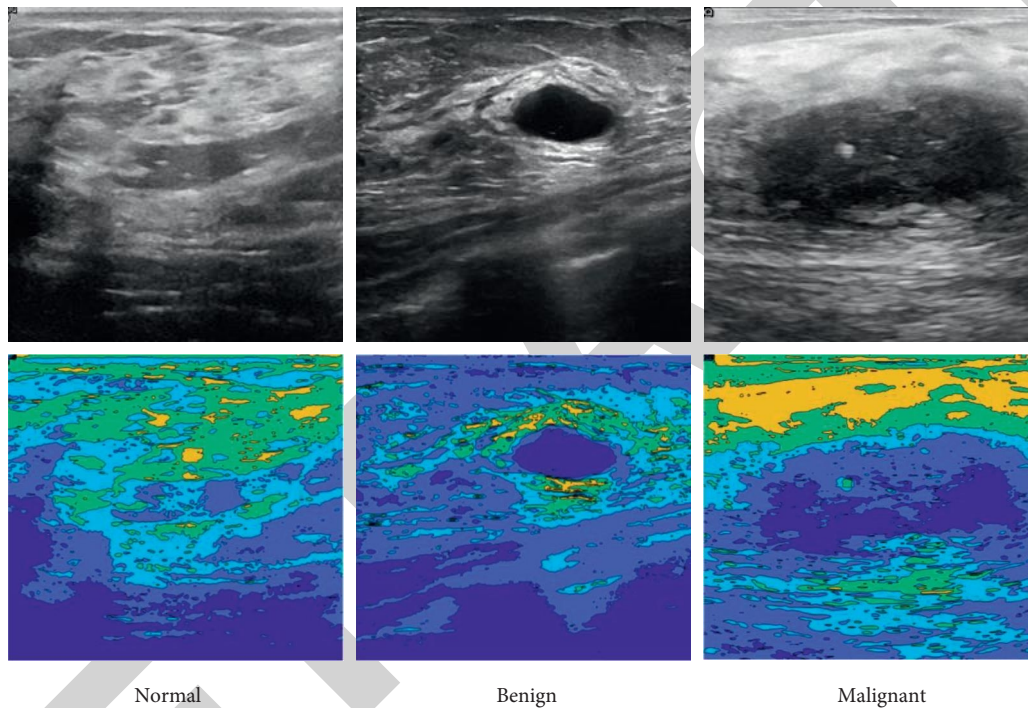


FIGURE 3: Example of ultrasound breast cancer images from the dataset.

specificity, precision, recall, and decreased data dimensionality and computing complexity [37].

Artificial neural networks and convolutional neural networks (CNN) are similar. They are composed of neurons with trainable connection weights. Each neuron takes some inputs, does a dot product, and then executes a nonlinearity if desired [38]. There is still a single variational scoring system from the raw picture pixels to class scores at the other end of the network. Furthermore, they still contain a loss function on the last (fully connected) layer (e.g., softmax). All of the learning strategies we devised for ordinary neural networks are still applicable. CNNs are valid for recognizing objects, people, and sceneries by looking for patterns in pictures. They can also categorize nonimage data, including audio, time series, and signaling data pretty well [39]. The confusion matrix is an accurately named instrument that best describes the classifier's performance. Knowing the confusion matrix necessitates learning a few definitions [40].

However, before we get into the concepts, let us look at a fundamental confusion matrix for binary or binomial identification with two categories (say, Y or N). Sensitivity refers to a classifier's capacity to choose all of the examples that must be selected. A perfect classifier will choose all true Ys and would not leave any true Ys out. To put it another way, there will be no false negatives. Any classification will miss many true Ys, resulting in false negatives. The capacity of the predictor to pick all instances that need to be chosen and refuse all cases that need to be denied is described as accurate [41].

4. Results

4.1. Data Collection. Breast ultrasound scans of women between 25 and 75 were first obtained. This information was compiled in 2018. There are 400 women among the sick. There are 780 images in the data collection, with an average

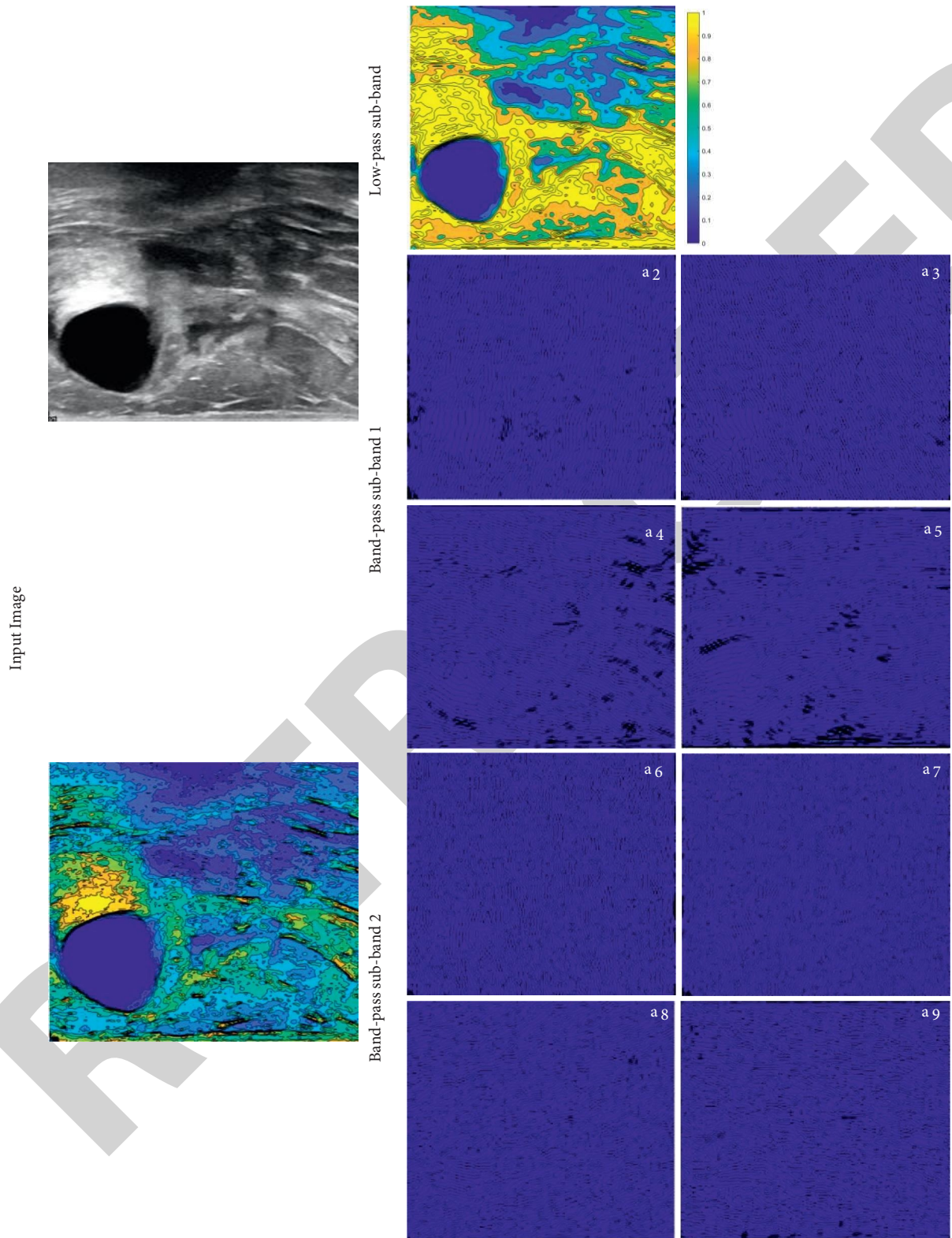


FIGURE 4: Results of contourlet decomposition.

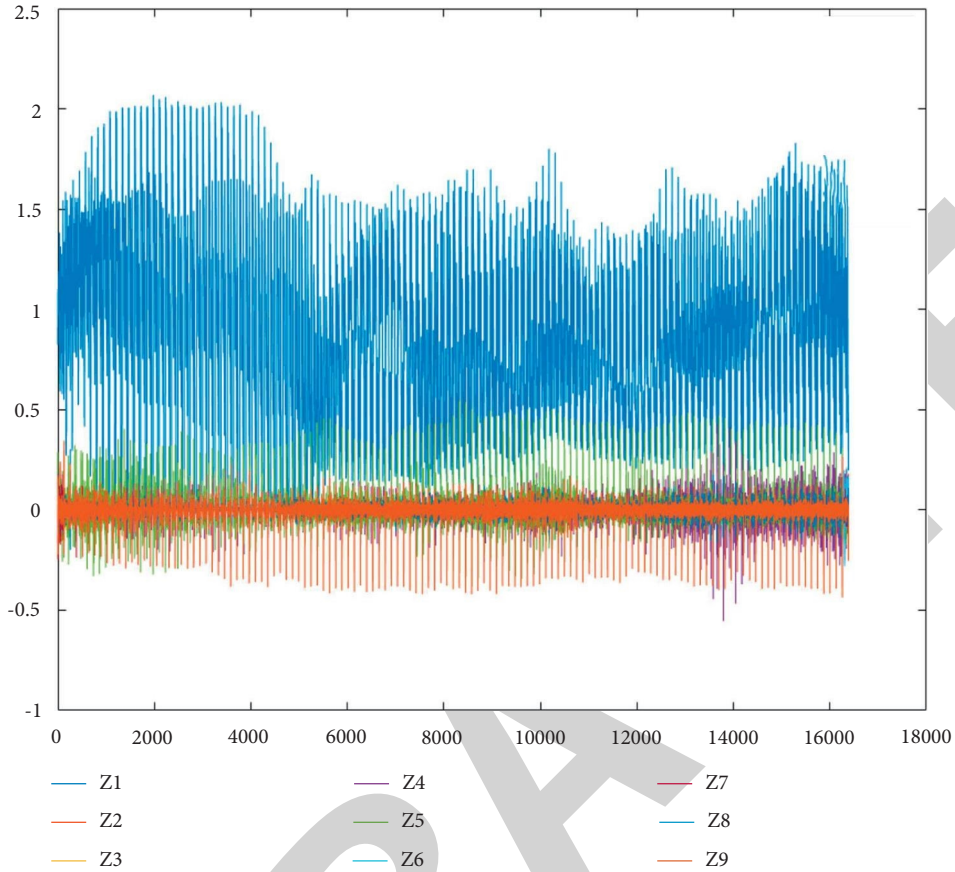


FIGURE 5: Reshaped sub-bands of contourlet transformation as feature signals.

```

Input:  $I_{m \times m} = \{m \times m\} \in R^2$ 
Output: Labels (Normal, Benign, Malignant)
Filter = pkva,
Number of levels =  $[2 \times 2]$ 
Smooth function = Raised cosine
Contourlet Decomposition:
 $[a_1, \dots, a_9] = \text{Contourlet}(I, \text{Filter}, \text{Number of levels}, \text{Smooth unction})$ 
Reshape sub-bands:
For  $i = 1:9$ 
    Sub_band_vector $_i = \text{Reshape}(a_i)$ 
End For
Feature Extraction:
For  $j = 1:9$ 
    For  $k = 1:7$ 
         $f_k^{(j)} = TD(\text{Sub\_band\_vector}_j)$ 
    End For
End For
 $F_{63 \times 1} = f_k^{(j)}, \quad j = 1, 2, \dots, 9; k = 1, 2, \dots, 7$ 
Feature reduction:
 $F' = \text{PCA}(F)$ 
Classification:
Train (KNN ( $F'$ ), Labels)
Train (SVM ( $F'$ ), Labels)
Train (LDA ( $F'$ ), Labels)
Train (DT ( $F'$ ), Labels)
Train (CNN ( $F'$ ), Labels)
Performance Analysis:
Plot (Confusion Matrix);
Plot (Performance Plot)
Plot (ROC)

```

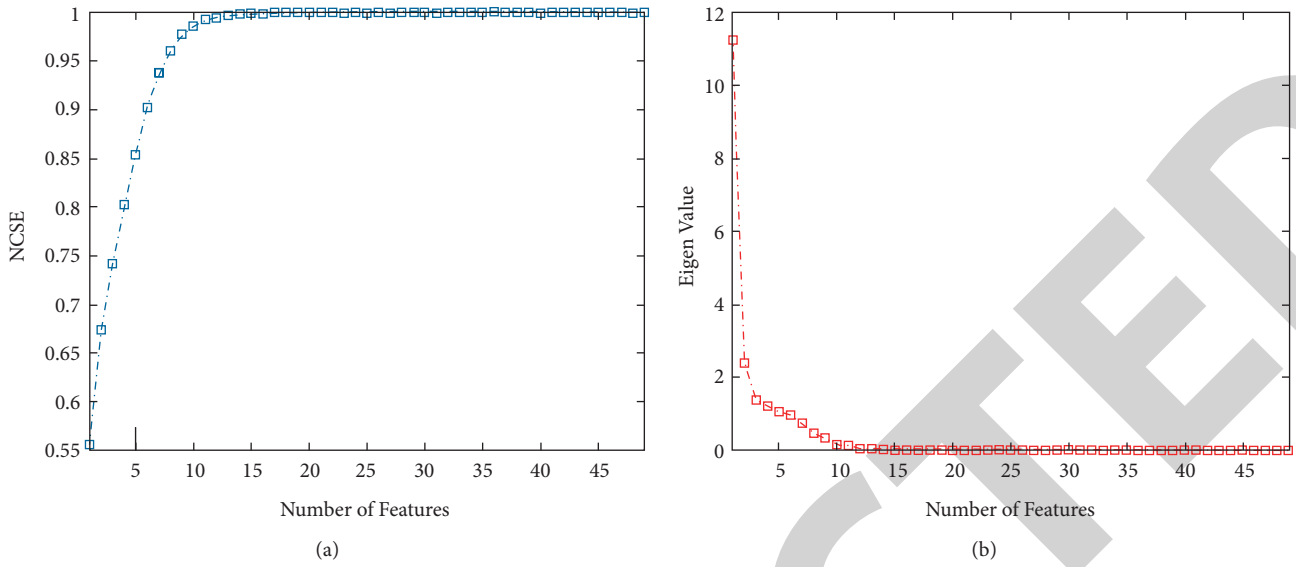


FIGURE 6: Scree plot and the (a) normalized cumulative; (b) summation of the eigenvalue (NCSE).

500 by 500 pixels size. The images are saved as PNG files. Normal, benign, and malignant images are divided into three categories. To decrease the processing complexity in this study, the picture size was reduced to 256 by 256 pixels for image classification and segmentation. Figure 3 shows an illustration of the images. Furthermore, the image's contour is shown to further demonstrate the data values in each class. The database is available online on [42].

4.2. Results of Contourlet Transformation. This study developed a unique feature extraction approach based on a contourlet transformation and time-dependent model combination. Each pyramid level has two directional filters in this decomposition contourlet transformation vector of numbers of directional filter bank decomposition levels at each pyramidal level. In addition, the number of tiers is regarded to be two. Figure 4 depicts the sub-bands of the proposed technique. One of the original images of the benign breast tumor is shown in Figure 4 (left). The images are displayed with contour to show the image and each sub-band. In each level, the transformation is carried out for two pyramid layers. At the first level of decomposition, the low-pass sub-band is not downsampled by the decomposition modes. The raised cosine function is the function handle for generating the pyramid decomposition filter:

$$\theta = 0.5 (1 - \cos(\pi x)),$$

$$\text{s.t.} \begin{cases} \theta(x \leq 0) = 0 \\ \theta(x \geq 1) = 1 \end{cases} \quad (4)$$

Moreover, the filter for the directional decomposition step is PKVA filtration. The resulted sub-bands were then used in the time-dependent model. Based on the results, the low-pass sub-band shows the tumor place. However, other sub-bands shows hidden parameter of the images. Evaluating the

correlation between each sub-bands and tumor type will be illustrated in the following sections. Regarding Figures 4 and 3, in the normal condition, there is no circular dark area to show the tumor place; however, in the benign image, the tumor place is the darkest circle of the image. On the other hand, a tumor in a malignant place is shown as a separate area. The contourlet transformation can illustrate the tumor with sub-bands to better diagnose the tumor.

4.3. Feature Extraction and Reduction. In this section, the results of the feature extraction are explained. The outcome of the contourlet transformation is nine sub-bands based on Figure 4. In the next step, after the decomposition of each image to sub-bands, the output matrices reshape to vector form. Therefore, each sub-bands shown participate in the feature extraction as a vector or pseudo-time series or signals. The reshaped signal of an image is depicted in Figure 5. Except for the low-pass signal, other vectors are oscillating over zero. Based on Table 3 and Algorithm 1, each sub-band extract seven features. Therefore, each input image creates $7 \times 9 = 63$ features.

Moreover, to reduce the computation time of classification, the feature vector dimension is reduced using PCA. This section used the normalized cumulative sum of eigenvalues (NCSE) to show new features' eigenvalue. The first ten features can satisfy the classification results based on the results. The feature reduction plots in Figure 6 are used to determine the best number of groups. We utilized the contourlet transformation system and time-dependent models in this diagram. The findings of the PCA approach indicate that we can identify images using ten features. The number of features was reduced from 63 to 10 by using PCA. Using the sub-bands of a contourlet transformation system for classification with fewer features can accelerate the classification method and improve accuracy.

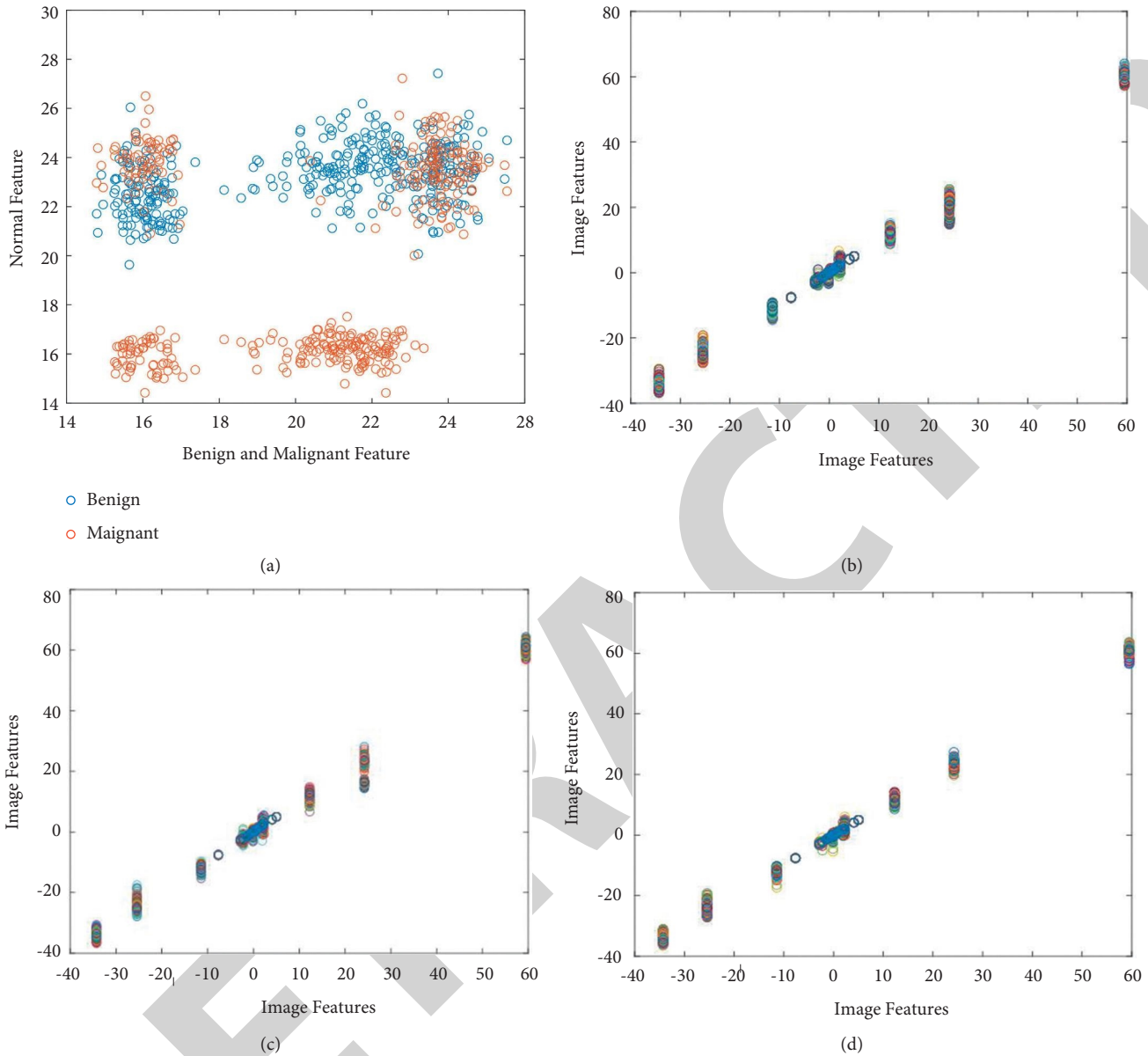
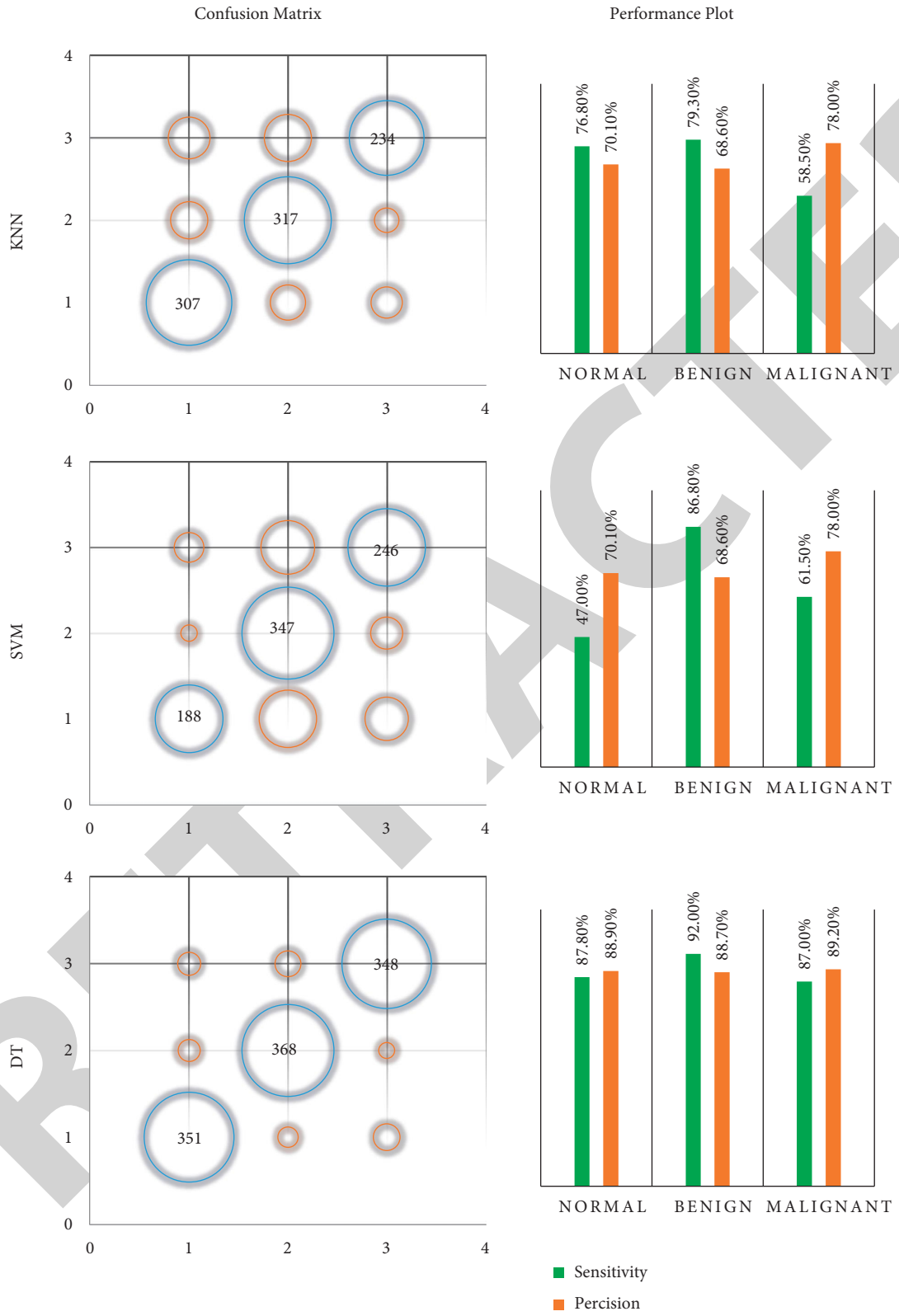


FIGURE 7: A scatter plot of breast cancer features according to their classes and the relationship between features, (a) Scatter, (b) normal, (c) malignant, and (d) benign.

To verify the features of contourlet transformation for classification, we studied the relationship between features. Figure 7(a) shows the scatter plot of benign and malignant features versus normal feature. Based on this plot, there is no direct relationship between the feature of each class. In other words, feature value in each class has no relationship, and each class's behavior is different. On the other hand, in Figures 7(b)–7(d), the normal, malignant, and benign class are illustrated. Based on this figure, there is a relationship between each sub-bands of CT features. It means a direct relationship between each sub-bands in all the normal or (benign/malignant). These facts verify that the utilized features can classify each class in a meaningful manner.

4.4. Classification Results. In this section, the classification is made using different machine learning methods. The input layer of the classification methods is ten reduced features of the images, and the output layer is the three-class label of normal, benign, and malignant. Total 1200 ultrasound images are used for the classification of breast cancer. The confusion matrices of the presented methods are illustrated in Figure 8. The blue balls show the true values, and the red balls are the false value of the classification. Moreover, labels 1, 2, and 3 show the normal, benign, and malignant, respectively. Regarding the results of the KNN method, from 400 input normal, benign, and malignant images, 307, 317, and 234 are detected correctly. Based on the results, the sensitivity of the KNN for diagnosing breast cancer for



(a)

FIGURE 8: Continued.

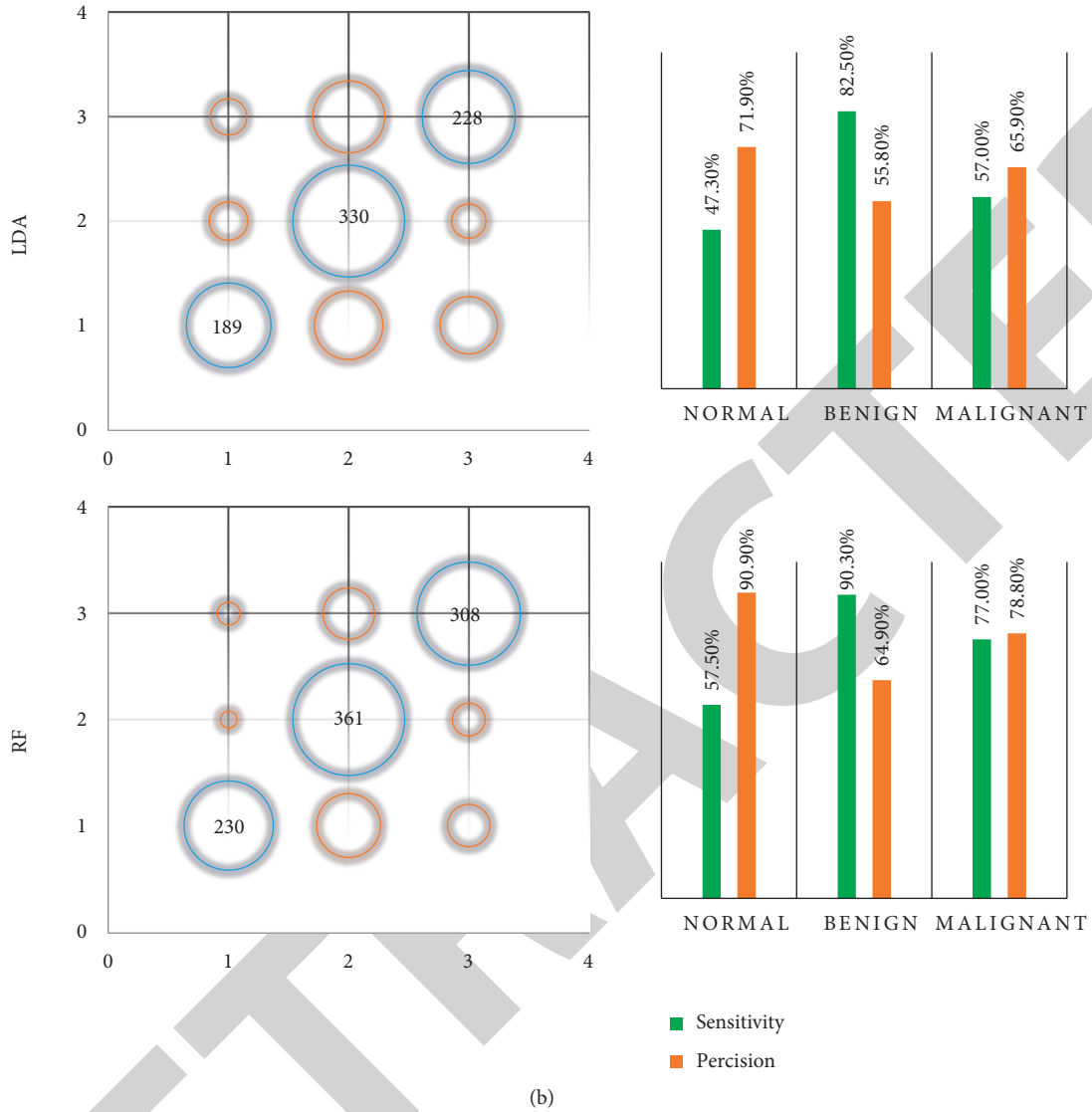


FIGURE 8: Confusion matrices and performance plot of the classification methods.

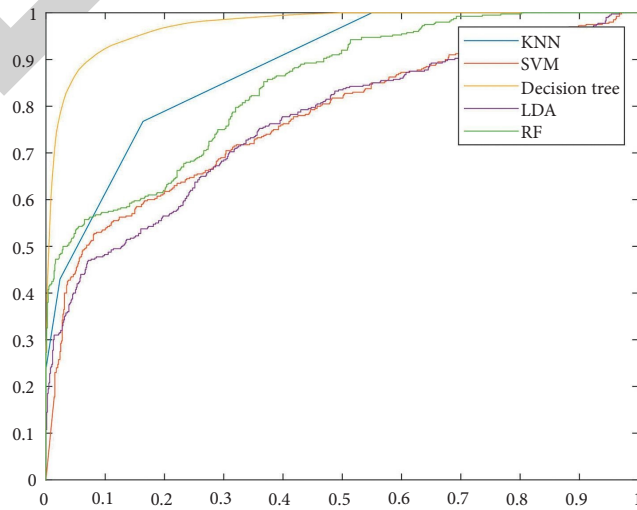


FIGURE 9: The ROC curve of the classifiers based on the presented feature extraction method.

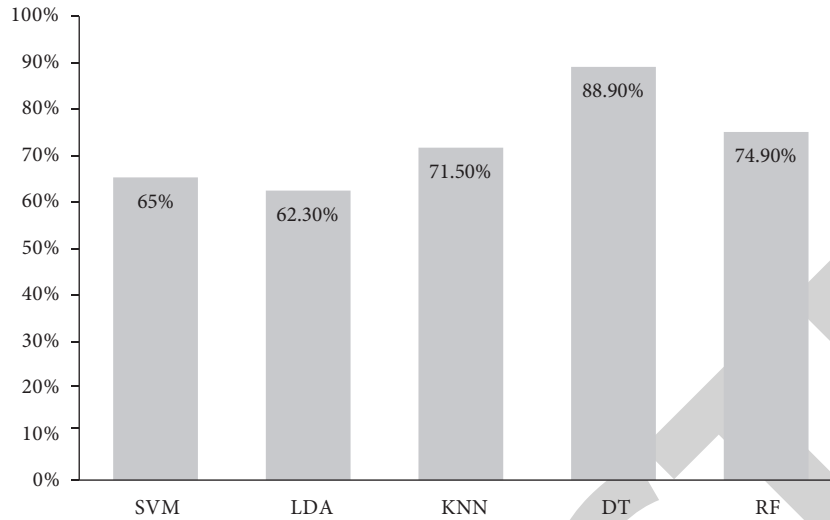


FIGURE 10: Accuracy value of the ML classifiers.

normal and benign is acceptable. Based on the results, the SVM and LDA approach reached the weak result for breast cancer diagnosis. However, the results of DT show that the sensitivity of the method is 87.8%, 92.0%, and 87.0%, respectively. The presented feature extraction method is compatible with the DT approach for this problem. In other words, 351, 368, and 348 ultrasound images from normal, benign, and malignant are detected, respectively. Moreover, the method's precision is 88.9%, 88.7%, and 89.2% for normal, benign, and malignant. Moreover, the RF results are also acceptable for diagnosing benign tumors with 90.3% sensitivity.

5. Discussion

To compare the presented machine learning method for diagnosing breast cancer, the ROC is depicted in Figure 9. Based on the ROC curve, the horizontal axis is the rate of the false-positive index based on the normal class. The vertical axis represents the true positive rate. The best classifier shows the highest true positive and lowest false-positive rates. The DT method shows the best classifier for the presented features based on the results. The accuracy of the machine learning classifiers is presented in Figure 10. Based on the results, SVM, LDA, KNN, DT, and RF accuracy is 65%, 62.3%, 71.5%, 88.90%, and 74.90%, respectively. Based on this chart, the DT architecture with the highest accuracy is the more accurate and compatible method for diagnosing breast cancer using the presented hybrid approach. Based on the literature review, the high diagnosis volume was performed using mammographic images with high accuracy. However, ultrasound image is more complex than mammography images, and designing a proper feature extraction method is essential. Therefore, this article presented a novel hybrid approach for extracting meaningful features to diagnose breast cancer. Based on the results, the presented method is acceptable for the classification of breast cancer with ultrasound images.

6. Conclusion

This study developed a unique feature extraction approach based on a contourlet transformation and time-dependent model combination. Each pyramid level has two directional filters in this decomposition contourlet transformation vector of numbers of directional filter bank decomposition levels at each pyramidal level. The sub-bands that emerged were then employed in the time-dependent model. The low-pass sub-band displays the tumor location; however, additional sub-bands reveal hidden visual parameters. After decomposing each picture into sub-bands, the resultant matrices are reshaped into vector form in the next phase. As a result, each sub-band displayed feature extraction as a vector, pseudo-time series, or signal. Each sub-band retrieved seven characteristics based on the results. As a result, each input image generates 63 distinct characteristics. Furthermore, the feature vector dimension is lowered using PCA to minimize classification calculation time. We examined the link between characteristics to validate the contourlet transformation features for categorization. According to the data, there is no direct association between the features of any class. To put it another way, the feature value of each class has no connection, and each class behaves differently. On the other hand, there is a link between the characteristics of each contourlet transformation sub-band. It indicates that in all normal or (benign/malignant) sub-bands, there is a direct link between them. These facts demonstrate that the traits used to classify each class are accurate. Different machine learning approaches are used to classify the data in this article. Breast cancer is classified using a total of 1200 ultrasound images. The DT findings reveal that the method's sensitivity is 87.8%, 92.0 percent, and 87.0 percent, respectively. It indicates that the feature extraction method is compatible with the DT approach to this problem. In other words, normal, benign, and malignant ultrasound pictures are discovered at 351, 368, and 348, respectively. Furthermore, the accuracy of the approach is 88.9%, 88.7%, and 89.2% for

normal, benign, and malignant, respectively. Based on the results, SVM, LDA, KNN, DT, and RF accuracy are 65 percent, 62.3 percent, 71.5 percent, 88.90 percent, and 74.90 percent, respectively. Using the provided hybrid methodology, the DT architecture with the highest accuracy is the most accurate and suitable way for diagnosing breast cancer.

Data Availability

Data are available and can be provided over the emails querying directly to the author at the corresponding author (soufia.bahmani@aut.ac.ir).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] F. F. Ting, Y. J. Tan, and K. S. Sim, "Convolutional neural network improvement for breast cancer classification," *Expert Systems with Applications*, vol. 120, pp. 103–115, 2019.
- [2] S. Sharma and S. Deshpande, "Breast cancer classification using machine learning algorithms," *Machine Learning for Predictive Analysis*, vol. 141, pp. 571–578, 2021.
- [3] M. Z. Alom, C. Yakopcic, M. S. Nasrin, T. M. Taha, and V. K. Asari, "Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network," *Journal of Digital Imaging*, vol. 32, no. 4, pp. 605–617, 2019.
- [4] N. K. Sakthivel, N. P. Gopalan, and S. Subasree, "G.-HWRP.: Gene signature based HRF cluster for predicting human diseases," *International Journal of Pure and Applied Mathematics*, vol. 117, no. 9, pp. 157–161, 2018.
- [5] M. Veta, J. P. W. Pluim, P. J. Van Diest, and M. A. Viergever, "Breast cancer histopathology image analysis: a review," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1400–1411, 2014.
- [6] J. G. Elmore, G. M. Longton, P. A. Carney et al., "Diagnostic concordance among pathologists interpreting breast biopsy specimens," *JAMA*, vol. 313, no. 11, pp. 1122–1132, 2015.
- [7] K. Dobruch-Sobczak, H. Piotrkowska-Wróblewska, Z. Klimonda et al., "Multiparametric ultrasound examination for response assessment in breast cancer patients undergoing neoadjuvant therapy," *Scientific Reports*, vol. 11, no. 1, pp. 1–9, 2021.
- [8] J. M. Chang, H. R. Koo, and W. K. Moon, "Radiologist-performed hand-held ultrasound screening at average risk of breast cancer: results from a single health screening center," *Acta Radiologica*, vol. 56, no. 6, pp. 652–658, 2015.
- [9] H. Moon, I. Jung, S. Park, M. Kim, J. Youk, and E. Kim, "Comparison of cancer yields and diagnostic performance of screening mammography vs. Supplemental screening ultrasound in 4394 women with average risk for breast cancer," *Ultraschall in der Medizin - European Journal of Ultrasound*, vol. 36, no. 03, pp. 255–263, 2015.
- [10] D. Muduli, R. Dash, and B. Majhi, "Automated breast cancer detection in digital mammograms: a moth flame optimization based ELM approach," *Biomedical Signal Processing and Control*, vol. 59, p. 101912, 2020.
- [11] J. G. Melekoodappattu and P. S. Subbian, "Automated breast cancer detection using hybrid extreme learning machine classifier," *Journal of Ambient Intelligence and Humanized Computing*, no. -10, p. 1, 2020.
- [12] S. Sasikala, M. Ezhilarasi, and S. Arun Kumar, "Detection of breast cancer using fusion of MLO and CC view features through a hybrid technique based on binary firefly algorithm and optimum-path forest classifier," in *Applied Nature-Inspired Computing: Algorithms and Case Studies*, pp. 23–40, Springer, Springer, Singapore, 2020.
- [13] S. S. Begum and D. R. Lakshmi, "Combining optimal wavelet statistical texture and recurrent neural network for tumour detection and classification over MRI," *Multimedia Tools and Applications*, vol. 79, no. 19-20, pp. 14009–14030, 2020.
- [14] X. Fei, S. Zhou, X. Han et al., "Doubly supervised parameter transfer classifier for diagnosis of breast cancer with imbalanced ultrasound imaging modalities," *Pattern Recognition*, vol. 120, p. 108139, 2021.
- [15] S. Yan, J. Yan, D. Liu et al., "A nano-predator of pathological MDMX construct by clearable supramolecular gold (I)-thiolpeptide complexes achieves safe and potent anti-tumor activity," *Theranostics*, vol. 11, no. 14, pp. 6833–6846, 2021.
- [16] Q. Xu, Y. Zeng, W. Tang et al., "Multi-task joint learning model for segmenting and classifying tongue images using a deep neural network," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2481–2489, 2020.
- [17] W. Tang, S. Wan, Z. Yang, A. E. Teschendorff, and Q. Zou, "Tumor origin detection with tissue-specific miRNA and DNA methylation markers," *Bioinformatics*, vol. 34, no. 3, pp. 398–406, 2018.
- [18] D. Q. Zeebaree, A. Mohsin Abdulazeez, D. Asaad Zebari, H. Haron, and H. Nuzly Abdull Hamed, "Multi-level fusion in ultrasound for cancer detection based on uniform LBP features," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 3363–3382, 2021.
- [19] G. Briganti, M. Scutari, and P. Linkowski, "A machine learning approach to relationships among alexithymia components," *Psychiatria Danubina*, vol. 32, no. 1, pp. 180–187, 2020.
- [20] A. Rezaei, H. Hassani, P. Moarefvand, and A. Golmohammadi, "Lithological mapping in Sangan region in Northeast Iran using ASTER satellite data and image processing methods," *Geology, Ecology, and Landscapes*, vol. 4, no. 1, pp. 59–70, 2020.
- [21] M. Zhang, Y. Chen, and W. Susilo, "PPO-CPQ: a privacy-preserving optimization of clinical pathway query for e-healthcare systems," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10660–10672, 2020.
- [22] H. Liu, J. Liu, S. Hou, T. Tao, and J. Han, "Perception consistency ultrasound image super-resolution via self-supervised CycleGAN," *Neural Computing and Applications*, pp. 1–11, 2021.
- [23] E. Eslami and H.-B. Yun, "Attention-based multi-scale convolutional neural network (A+ MCNN) for multi-class classification in road images," *Sensors*, vol. 21, no. 15, p. 5137, 2021.
- [24] E. Sadeghipour, N. Sahragard, M.-R. Sayebani, and R. Mahdizadeh, "Breast cancer detection based on a hybrid approach of firefly algorithm and intelligent systems," *Indian Journal of Fundamental and Applied Life Sciences*, vol. 5, p. S1, 2015.
- [25] M. Rezaei, F. Farahanipad, A. Dillhoff, R. Elmasri, and V. Athitsos, "Weakly-supervised hand part segmentation from depth images," in *Proceedings of the 14th Pervasive Technologies Related to Assistive Environments Conference*, pp. 218–225, Corfu, Greece, 2021.

Retraction

Retracted: Automatic Target Recognition of SAR Images Using Collaborative Representation

Computational Intelligence and Neuroscience

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

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

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

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

References

- [1] J. Hu, "Automatic Target Recognition of SAR Images Using Collaborative Representation," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3100028, 7 pages, 2022.

Research Article

Automatic Target Recognition of SAR Images Using Collaborative Representation

Jinge Hu 

Chongqing Three Gorges University, Chongqing 404100, China

Correspondence should be addressed to Jinge Hu; 20171002@sanxiau.edu.cn

Received 2 April 2022; Accepted 4 May 2022; Published 24 May 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Jinge 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.

Synthetic aperture radar (SAR) automatic target recognition (ATR) is one of the key technologies for SAR image interpretation. This paper proposes a SAR target recognition method based on collaborative representation-based classification (CRC). The collaborative coding adopts the global dictionary constructed by training samples of all categories to optimally reconstruct the test samples and determines the target category according to the reconstruction error of each category. Compared with the sparse representation methods, the collaborative representation strategy can improve the representation ability of a small number of training samples for test samples. For SAR target recognition, the resources of training samples are very limited. Therefore, the collaborative representation is more suitable. Based on the MSTAR dataset, the experiments are carried out under a variety of conditions and the proposed method is compared with other classifiers. Experimental results show that the proposed method can achieve superior recognition performance under the standard operating condition (SOC), configuration variances, depression angle variances, and a small number of training samples, which proves its effectiveness.

1. Introduction

Synthetic aperture radar (SAR) can work continuously under all-weather conditions, thus providing an effective tool for intelligence reconnaissance. SAR automatic target recognition (ATR), as an important research content of SAR image interpretation [1], has been extensively studied in the past 20 years. Specific SAR target recognition methods mainly focus on two technologies: feature extraction and classifier design. Feature extraction aims to reduce the dimensionality of the original SAR data, thereby facilitating the subsequent decision making. The features commonly used in SAR target recognition include target region, target contour [2–7], scattering centers [8–10], principal component analysis (PCA) features, linear discriminant analysis (LDA) features, image decompositions [11–18], etc. The classifier design aims to make decisions on the original SAR image or the extracted features to determine the target category. With the advancement of pattern recognition technology, a large number of classifiers have been applied in SAR target recognition, such as K-nearest neighbor (KNN)

[8], support vector machine (SVM) [19–21], sparse representation-based classification (SRC) [22–27], convolutional neural network (CNN) [28–39], and so on. Mishra applied PCA and LDA to SAR image feature extraction and classified them through KNN [11]. The authors in [2] first extracted the elliptical Fourier descriptor of the SAR target contour and then used the SVM classifier for target recognition. In [8], the authors proposed an attribute scattering center matching algorithm and applied it to SAR target recognition. The authors in [19, 23] applied SVM and SRC to SAR target recognition, respectively, and proved their effectiveness. The scope of application and classification performance of different classifiers are usually not the same. Therefore, in order to improve the recognition performance, it is necessary to select a more effective and robust classifier.

In this paper, the collaborative representation-based classification (CRC) [40–42] is applied to SAR target recognition. The basic idea is to use a global dictionary composed of various training samples to optimally reconstruct test samples and then determine the target category according to various reconstruction errors. Zhang et al. [40]

first proposed a collaborative representation classifier and applied it to face recognition with good results. The results also proved that under the condition of limited training samples in each category, the strategy of using collaborative representation is more robust than the traditional sparse representation. For SAR target recognition, limited by the ability to acquire data, the number of training samples in each category is often very limited, so it is more appropriate to use collaborative representation during the classification. In order to scientifically test the proposed method, the moving and stationary target acquisition and recognition (MSTAR) dataset is used to conduct verification experiments under different conditions, and the superiority of the proposed method is verified by comparison with other classifiers.

2. SRC

Sparse representation is essentially a linear representation theory [22–27]. Different from the general linear representation, the sparse representation requires the linear representation coefficients to be sparse, that is, only a small number of coefficients are nonzero. Suppose that the global dictionary constructed by C classes of targets is $A = [A^1, A^2, \dots, A^C] \in R^{d \times N}$, where d is the dimension of each atom (i.e., the feature vector extracted from the training sample) and $A^i \in R^{d \times N_i}$ ($i = 1, 2, \dots, C$) denotes the N_i atoms from the i th class. For the test sample $y \in R^d$, the following equation is used for sparse representation:

$$\hat{\alpha} = \arg \min \|\alpha\|_0 \text{ s.t. } \|y - A\alpha\|_2 \leq \varepsilon. \quad (1)$$

In (1), α represents the sparse representation coefficient vector and ε is the reconstruction error. The above problem is difficult to be directly solved because it involves the ℓ_0 norm optimization. According to the research results of compressive sensing theory, the ℓ_1 minimization algorithm, orthogonal matching pursuit (OMP), Bayesian compressive sensing (BCS), etc. can be used to obtain the approximate solution of problem (1).

On the basis of obtaining the sparse representation coefficient vector $\hat{\alpha}$, the target category is determined by calculating the reconstruction error of each training category with regard to the test sample, as shown below:

$$\begin{aligned} r(i) &= \|y - A_i \hat{\alpha}_i\|_2^2 \quad (i = 1, 2, \dots, C), \\ \text{identity}(y) &= \arg \min_i (r(i)). \end{aligned} \quad (2)$$

In (2), $\hat{\alpha}_i$ represents the partial coefficients corresponding to the i th class and $r(i)$ ($i = 1, 2, \dots, C$) is the reconstruction error of each training class to the test sample. Finally, the test sample is determined as the category that obtains the smallest reconstruction error.

3. CRC

The sparse representation in SRC emphasizes the effect of sparsity on target recognition but ignores the complementarity of different categories. When the number of training samples of each category in the dictionary is limited,

it is difficult to achieve a high-precision description of the test samples by using sparse representation. In this situation, the classification results based on sparse representation may bring false judgments. Collaborative representation means that the joint use of samples of different types can provide a more accurate description of the test samples. Therefore, CRC uses the least regularized mean square error algorithm to solve the collaborative coding coefficients on the global dictionary as follows [39]:

$$\hat{x} = \arg \min \{\|y - Ax\|_2^2 + \lambda \|x\|_2^2\}. \quad (3)$$

In (3), λ is the regularization coefficient. Compared with the sparse representation coefficients, the solution of cooperative coding is much simpler. The analytical solution of (3) can be calculated as follows:

$$\hat{x} = (A^T A + \lambda * I)^{-1} A^T y. \quad (4)$$

In (4), I is an identity matrix. According to the solved collaborative representation coefficient, using the same method as (2), the reconstruction error of each category is calculated and the target category is judged based on the smallest representation error. Compared with SRC, the collaborative representation strategy can better characterize the test samples. In addition, the ℓ_2 norm minimization in (3) is a convex optimization algorithm, which is simpler and more efficient than the ℓ_0 norm optimization in (1).

This paper applies CRC to SAR target recognition, and its specific implementation is described as follows. In order to reduce the dimensionality of the original SAR data, PCA is used for feature extraction. The PCA features of the training samples are used to construct a global dictionary, and then the PCA feature of the test sample is collaborative represented based on the global dictionary. Finally, the target category is determined according to the resulting reconstruction errors.

4. Experiments

4.1. Description of Dataset. In order to verify the recognition performance of the method in this paper, the experiments are carried out on the MSTAR dataset. The dataset collects the measured SAR images of 10 types of ground stationary targets by the X-band sensors, which were published by the US DARPA/AFRL. The image resolution is $0.3 \text{ m} \times 0.3 \text{ m}$, which is an important dataset for testing SAR target recognition methods. Figure 1 shows some examples of the optical and SAR images of 10 types of targets. Table 1 lists the training and test samples used in this paper. In the experiment, 10 types of target images at the depression angle of 17° are used as training samples, and 10 types of target images at the depression angle of 15° are used as the test samples to be recognized.

4.2. Results and Discussion. In order to fully verify the effectiveness of the method in this paper, the CRC is compared with several classic SAR target recognition classifiers, including the KNN, SVM, SRC, and CNN. In order to ensure the consistency of the comparison, the classifiers KNN,

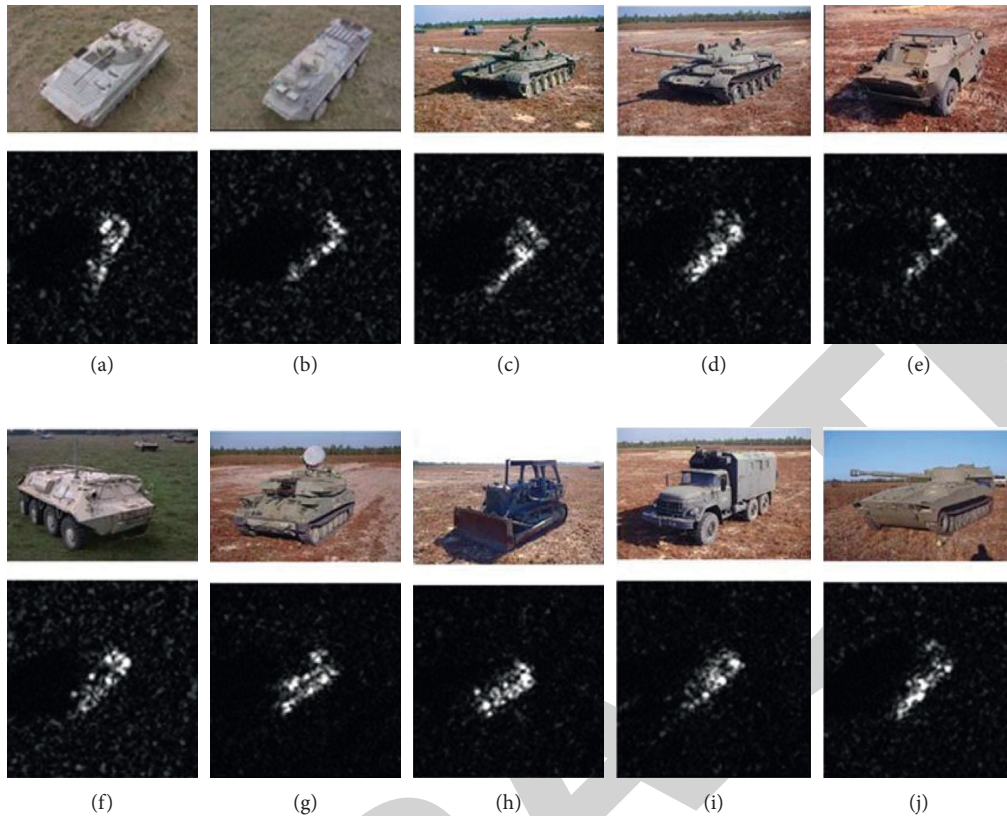


FIGURE 1: Optical and SAR images of the targets in the MSTAR dataset. (a) BMP2, (b) BTR70, (c) T72, (d) T62, (e) BRDM2, (f) BTR60, (g) ZSU23/4, (h) D7, (i) ZIL131, (j) 2S1.

TABLE 1: Training and test samples used in this paper.

Type	BMP2	BTR70	T72	T62	BDRM2	BTR60	ZSU23/4	D7	ZIL131	2S1
Training set	233 (Sn_9563)	233	232 (Sn_132)	299	298	256	299	299	299	299
Test set	195 (Sn_9563) 196 (Sn_9566) 196 (Sn_c21)	196	196 (Sn_132) 195 (Sn_812) 191 (Sn_s7)	273	274	195	274	274	274	274

SVM, and SRC also classify the extracted PCA feature vectors like CRC. All methods are run based on MATLAB 2016, and the hardware platform is a personal computer with a 3.4 GHz Intel i7 processor and 8G memory.

4.2.1. 10-Class Recognition. First, the test samples of 10 types of targets in Table 1 are classified based on the training samples. The recognition results of the method in this paper are shown in Figure 2. It can be seen that the CRC method proposed in this paper has an average recognition rate of 96.92% for 10 types of targets, which fully demonstrates its effectiveness. Table 2 compares the average recognition rates and time consumption of CRC and other types of SAR target recognition methods for 10 types of targets. The method in this paper has the highest recognition rate, which shows the high performance of the proposed method. By comparing the CRC and SRC methods, it can be seen that the collaborative representation strategy adopted by CRC effectively improves the target recognition performance compared to the sparse

representation algorithm. Comparing the time consumption of various methods to recognize a single MSTAR image under the same hardware platform, the proposed method has higher efficiency. As shown in (4), the CRC solution can be smoothly obtained using the convex optimization. Therefore, it can achieve much higher efficiency. The performance of the CNN method ranks second under this condition mainly because of the high classification capability of deep learning when the training samples are sufficient to cover the test samples.

4.2.2. Performance at Different Feature Dimensions. The performance of the classifier is closely related to the extracted features. In order to further test the recognition robustness of the CRC method, this paper conducts recognition experiments on 10 types of targets in different feature dimensions. Table 3 shows the performance curves of the KNN, SVM, and SRC methods as the PCA feature dimensions change. All these methods maintain a relatively close recognition performance within a given feature

BMP2	0.957	0.005	0.019	0.002	0.007	0.000	0.000	0.003	0.002	0.005
BTR70	0.000	0.990	0.000	0.005	0.000	0.005	0.000	0.000	0.000	0.000
T72	0.015	0.000	0.964	0.007	0.003	0.005	0.000	0.000	0.003	0.002
T62	0.004	0.000	0.000	0.989	0.000	0.004	0.004	0.000	0.000	0.000
BDRM2	0.000	0.004	0.000	0.000	0.989	0.000	0.000	0.004	0.004	0.000
BTR60	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
ZSU23/4	0.000	0.000	0.000	0.004	0.000	0.000	0.996	0.000	0.000	0.000
D7	0.004	0.000	0.004	0.007	0.000	0.000	0.000	0.985	0.000	0.000
ZIL131	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.996	0.000
2S1	0.004	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.993
	BMP2	BTR70	T72	T62	BDRM2	BTR60	ZSU23/4	D7	ZIL131	2S1

FIGURE 2: The recognition results of 10 classes of targets.

TABLE 2: Performances of different methods on 10 classes of targets.

Method type	CRC	KNN	SRC	SVM	CNN
Average recognition rate (%)	97.22	94.35	95.64	95.82	96.04
Time consumption (ms)	25.3	40.3	32.1	24.6	36.23

TABLE 3: The recognition performances of different methods at different feature dimensions.

Feature dimension	40	60	80	100	120
Average recognition rate (%)	95.23	96.02	96.92	96.38	96.12

dimension. In contrast, the CRC method proposed in this paper has the highest average recognition rate in each dimension, which proves its superiority. The CNN method is not compared in this condition because the deep networks are directly trained based on the image intensities rather than the extracted PCA feature vectors.

4.2.3. Configuration Variance. Configuration variances are more common in military targets. For example, a certain type of tank can be transformed into another configuration through local structural adjustments. This type of difference will have a certain impact on target recognition performance. In order to test the recognition performance of the CRC method under configuration variances, part of the training and test samples in Table 1 is used to perform recognition experiments, as shown in Table 4. Among them, the training and test samples of BMP2 and T72 targets have some configuration differences. Table 5 lists the recognition performance of various methods under configuration variances. Compared with the average recognition rates under the 10-class recognition condition in Table 2, the performance of

TABLE 4: Training and test samples of configuration variance.

Target type	Training set	Test set
BMP2	233 (Sn_9563)	196 (Sn_9566) 196 (Sn_c21)
BTR70	233 (Sn_c71)	196 (Sn_c71)
T72	232 (Sn_132)	195 (Sn_812) 191 (Sn_s7)

TABLE 5: Comparison with other methods under configuration variance.

Method type	CRC	KNN	SRC	SVM	CNN
Average recognition rate (%)	95.53	92.32	94.15	93.98	94.28

various methods under the condition of configuration variances has declined to varying degrees. In contrast, the CRC method proposed in this paper achieves the highest recognition performance under configuration variances, reflecting its effectiveness for configuration changes. Collaborative representation has stronger representation ability for test samples with different configurations, thereby improving the performance of target recognition.

4.2.4. Depression Angle Variance. The difference in the depression angle will cause the acquired SAR image to undergo a large change. At this time, the performance of target recognition tends to be severely degraded. Table 6 shows the training and test samples used in this experiment. There is a big difference in depression angle between the training and the test sets. The recognition performance of various algorithms is shown in Table 7. When the change in depression angle is relatively small (e.g., from 17° to 30°), various methods can still have higher recognition rates. When the

TABLE 6: Training and test sets from different depression angles.

Class	Training		Test	
	Depression	Number of samples	Depression	Number of samples
2S1	17°	299	30°	288
			45°	303
BDRM2	17°	298	30° 303	287
ZSU23/4	17°	299	30°	288
			45°	303

TABLE 7: Performances of different methods under large depression angle variance.

Method type	Average recognition rate (%)	
	30°	45°
CRC	97.96	72.65
KNN	95.12	64.48
SRC	96.15	69.12
SVM	95.24	67.56
CNN	96.38	68.09

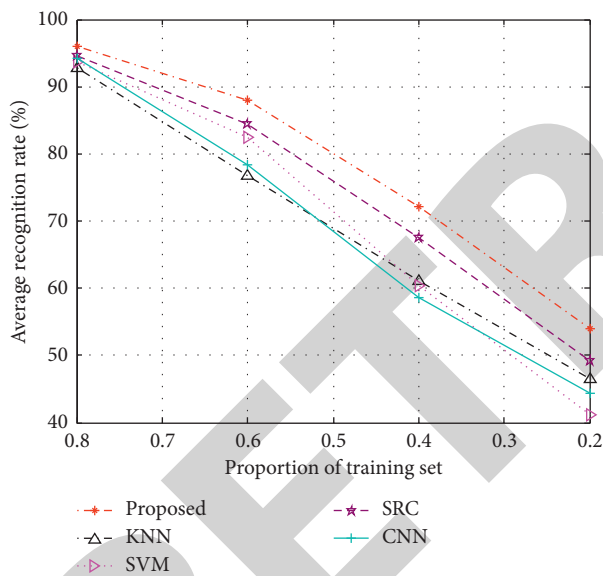


FIGURE 3: Comparison of results under few training samples.

depression angle changes greatly (e.g., from 17° to 45°), the recognition performance of various algorithms has experienced a sharp drop. In contrast, the CRC method proposed in this paper maintains the highest average recognition rates at both 30° and 45° depression angles, which proves that it has strong robustness to depression angle changes. Similar to the case of configuration variances, the collaborative representation has good capability to represent the test samples with large depression angle differences. Therefore, the CRC method proposed in this paper could achieve better robustness to possible depression angle variances.

4.2.5. Few Training Samples. Limited by the ability of SAR data acquisition, the training samples available in the actual

process are very limited compared to the large number of samples to be identified. Therefore, the stability of the recognition algorithm under a small number of training samples is very critical. In order to test the robustness of the CRC method under a small number of training samples, this paper takes 20%, 40%, 60%, and 80% of the original training samples in Table 1 to form the training set and then test the recognition performance of different methods. The performance comparison of various methods is shown in Figure 3. It can be seen that with the reduction of training samples, the recognition performance of various methods has been significantly reduced. In contrast, the CRC method is least affected by the reduction of training samples and maintains the highest average recognition rates on the training sets of various sizes. Therefore, the collaborative representation strategy adopted by CRC effectively improves the recognition performance of the proposed method under a small number of training samples.

5. Conclusion

This paper proposes a SAR target recognition method based on CRC. Compared with the traditional sparse representation, the collaborative representation strategy can comprehensively utilize various training samples to achieve higher-precision representations of the test samples, thereby improving the ability to represent the test samples under the condition of a small number of training samples. The CRC is applied to SAR target recognition, and experimental verification is carried out on the MSTAR dataset. The results show that the method in this paper has a recognition rate of 96.92% for 10 types of MSTAR targets, which shows its good performance. Through testing the recognition performance under different feature dimensions, configuration variances, depression angle variances, and a small number of training samples, the results further show the robustness of the CRC method to different kinds of extended operating conditions. All these illustrate the superiority of the method in this paper and prove that it has great application potential in SAR target recognition.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was supported by Foundation of Chongqing Municipal Key Laboratory of Institutions of Higher Education ([2017]3), Foundation of Chongqing Development and Reform Commission (2017[1007]), and Foundation of Chongqing Three Gorges University.

References

- [1] K. El-Darymli, E. W. Gill, P. McGuire, D. Power, and C. Moloney, "Automatic target recognition in synthetic aperture radar imagery: a state-of-the-art review," *IEEE Access*, vol. 4, pp. 6014–6058, 2016.
- [2] M. Amoon and G. Rezaei-rad, "Automatic target recognition of synthetic aperture radar (SAR) images based on optimal selection of Zernike moments features," *IET Computer Vision*, vol. 8, no. 2, pp. 77–85, 2014.
- [3] B. Ding, G. Wen, C. Ma, and X. Yang, "Target recognition in synthetic aperture radar images using binary morphological operations," *Journal of Applied Remote Sensing*, vol. 10, no. 4, Article ID 046006, 2016.
- [4] C. Shan, B. Huang, and M. Li, "Binary Morphological Filtering of Dominant Scattering Area Residues for SAR Target Recognition," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 9680465, 15 pages, 2018.
- [5] L. Jin, J. Chen, and X. Peng, "Synthetic aperture radar target classification via joint sparse representation of multi-level dominant scattering images," *Optik*, vol. 186, pp. 110–119, 2019.
- [6] J. Tan, X. Fan, S. Wang et al., "Target recognition of SAR images by partially matching of target outlines," *Journal of Electromagnetic Waves and Applications*, vol. 33, no. 7, pp. 865–881, 2019.
- [7] B. Feng, W. Tang, and D. Feng, "Target recognition of SAR images via hierarchical fusion of complementary features," *Optik*, vol. 217, Article ID 164695, 2020.
- [8] L. C. Potter and R. L. Mose, "Attributed scattering centers for SAR ATR," *IEEE Transactions on Image Processing*, vol. 6, no. 1, pp. 79–91, 1997.
- [9] B. Ding, G. Wen, J. Zhong, C. Ma, and X. Yang, "A robust similarity measure for attributed scattering center sets with application to SAR ATR," *Neurocomputing*, vol. 219, pp. 130–143, 2017.
- [10] B. Ding, G. Wen, X. Huang, C. Ma, and X. Yang, "Target recognition in synthetic aperture radar images via matching of attributed scattering centers," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 7, pp. 3334–3347, 2017.
- [11] A. K. Mishra, "Validation of PCA and LDA for SAR ATR," in *Proceedings of The in IEEE tencon 2008*, pp. 1–6, Hyderabad, India, November 2008.
- [12] Z. Cui, Z. Cao, J. Yang, J. Feng, and H. Ren, "Target recognition in synthetic aperture radar images via non-negative matrix factorisation," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1376–1385, 2015.
- [13] W. Xiong, L. Cao, and Z. Hao, "Combining wavelet invariant moments and relevance vector machine for SAR target recognition," in *Proceedings of the in iet international radar conference*, pp. 1–4, Guilin, April 2009.
- [14] G. Dong, G. Kuang, N. Wang, L. Zhao, and J. Lu, "SAR target recognition via joint sparse representation of monogenic signal," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 7, pp. 3316–3328, 2015.
- [15] Y. Zhou, Y. Chen, R. Gao, J. Feng, P. Zhao, and L. Wang, "SAR target recognition via joint sparse representation of monogenic components with 2D canonical correlation analysis," *IEEE Access*, vol. 7, p. 1, Article ID 25826, 2019.
- [16] M. Chang, X. You, and Z. Cao, "Bidimensional empirical mode decomposition for SAR image feature extraction with application to target recognition," *IEEE Access*, vol. 7, Article ID 135731, 2019.
- [17] M. Yu, G. Dong, H. Fan, and G. Kuang, "SAR target recognition via local sparse representation of multi-manifold regularized low-rank approximation," *Remote Sensing*, vol. 10, no. 2, p. 211, 2018.
- [18] Y. Huang, J. Peia, J. Yanga, B. Wang, and X. Liu, "Neighborhood geometric center scaling embedding for SAR ATR," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 1, pp. 180–192, 2014.
- [19] Q. Zhao and J. Principe, "Support vector machines for SAR automatic target recognition," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 2, pp. 643–654, 2001.
- [20] C. Tison, N. Pourthie, and J. Souyris, "Target Recognition in SAR Images with Support Vector Machines (SVM)," in *Proceeding of the 2007 IEEE international geoscience and remote sensing symposium*, pp. 456–459, Barcelona, Spain, July 2007.
- [21] M. E. Demirhan and Ö. Salor, "Classification of targets in SAR images using SVM and k-NN techniques," in *Proceedings of the 2016 24th signal processing and communication application conference (siu)*, pp. 1581–1584, Zonguldak, Turkey, May 2016.
- [22] H. Liu and S. Li, "Decision fusion of sparse representation and support vector machine for SAR image target recognition," *Neurocomputing*, vol. 113, pp. 97–104, 2013.
- [23] J. J. Thiagaraiann, K. N. Ramamurthy, P. Knee, A. Spanias, and V. Berisha, "Sparse representations for automatic target classification in SAR images," in *Proceedings of the in 4th international Symposium commun., control signal process*, pp. 1–4, Limassol, Cyprus, March 2010.
- [24] H. Song, K. Ji, and Y. Zhang, "Sparse representation-based SAR image target classification on the 10-class MSTAR data set," *Applied Sciences*, vol. 6, 2016.
- [25] B. Ding and G. Wen, "Sparsity constraint nearest subspace classifier for target recognition of SAR images," *Journal of Visual Communication and Image Representation*, vol. 52, pp. 170–176, 2018.
- [26] W. Li, J. Yang, and Y. Ma, "Target recognition of synthetic aperture radar images based on two-phase sparse representation," *Journal of Sensors*, vol. 2020, Article ID 2032645, 12 pages, 2020.
- [27] L. Yu, L. Wang, and Y. Xu, "Combination of Joint Representation and Adaptive Weighting for Multiple Features with Application to SAR Target Recognition," *Scientific Programming*, vol. 2021, Article ID 9063419, 9 pages, 2021.
- [28] X. X. Zhu, D. Tuia, L. Mou et al., "Deep learning in remote sensing: a comprehensive review and list of resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017.
- [29] M. Kang, K. Ji, X. Leng, X. Xing, and H. Zou, "Synthetic aperture radar target recognition with feature fusion based on a stacked autoencoder," *Sensors*, vol. 17, no. 12, p. 192, 2017.
- [30] D. A. E. Morgan, "Deep convolutional neural networks for ATR from SAR imagery," in *Proceedings of the in*, pp. 1–13, Baltimore, Maryland, USA, May 2015.

Research Article

Construction of Enterprise Financial Early Warning Model Based on Logistic Regression and BP Neural Network

Jincheng Lyu 

Tianjin Xiqing Economic Development Co., Ltd., Tianjin 300380, China

Correspondence should be addressed to Jincheng Lyu; 14125243@bjtu.edu.cn

Received 22 March 2022; Accepted 15 April 2022; Published 24 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Jincheng Lyu. 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.

At present, the number of enterprises in financial crisis in China is rising sharply, and the ability of enterprises to resist risks is generally weak. Therefore, it is necessary to establish a corporate financial crisis early warning system, to detect the signs of corporate financial crisis before it arrives and to inform managers in advance, so that effective measures can be taken as soon as possible to eliminate hidden dangers. This paper selects the two-year data of 40 companies from 2017 to 2019 as training samples and the data of 20 companies as prediction samples. After testing, 12 index variables that can reflect the financial problems of energy companies are finally selected as the basis for modeling. Then, we use Logistic and BP neural network modeling, respectively, to study and compare the data from 2017 to 2019 to predict the financial risk in the following year. The results show that the BP neural network model in the two models is better than the Logistic model in terms of fitting degree or prediction accuracy for enterprise financial early warning. Therefore, the BP neural network model has a better effect and is more suitable for the practical application of enterprises in China.

1. Introduction

With the rapid development of the economy in China, small- and medium-sized enterprises, as the most vital group in China, have grown into an important part of the economy, whether in foreign export trade or in increasing employment opportunities and increasing GDP. At present, China has 10 million industrial and commercial registered enterprises, of which 99% are small- and medium-sized enterprises. Small- and medium-sized enterprises provide 80% of the urban population's employment, and 70% of innovation comes from small- and medium-sized enterprises. In the competitive environment dominated by the buyer's market, the market competition continues to intensify [1–3]. In order to consolidate or expand their competitive advantages, small- and medium-sized enterprises in the industry must provide customers with competitive products and services, including preferential credit settlement methods, mainly in the form of credit sales. However, this method will make the enterprise itself bear the risk of the other party's use, which often causes the accounts receivable to be unable to be recovered in full

and on schedule, and many enterprises fall into the dilemma of “not selling on credit and waiting to die and selling on credit to court death.” This requires enterprises to fully understand their customers, develop a reasonable financial system, and lock down the financial risks of the enterprise by conducting financial early warning assessments of their finances before cooperating [4–7].

Although small- and medium-sized enterprises have made significant contributions to the economic growth in China, due to their lack of strict accounting systems, lack of professional financial personnel, and relatively small scale, it is currently difficult for small- and medium-sized enterprises to obtain financing with high financing costs to obtain financing from financial institutions. At this stage, banks, other enterprises, and financial institutions do not have a reasonable financial early warning system to assess the financial risks and experience of SMEs, and they will not easily lend to SMEs. The problem of shortage of funds has limited the development of enterprises to a certain extent [8–10].

Since the 1930s, many scholars have devoted a lot of time and energy to the study of financial crisis early warning and

have also developed some more objective, rigorous, and at the same time simple and feasible assessment methods. Fitzpatrick (1932) studied a sample of 19 firms and used a single financial ratio to divide the sample into bankrupt and nonbankrupt groups. It is found that net profit/shareholders' equity and shareholders' equity/liabilities have a good ability to discriminate financial crisis, and these two ratios show significant differences in the first three years of operating failure. Secrist (1938) used the ratio of assets/liabilities to compare the difference between failed banks and normal banks [11–13]. The research of Fitz et al. is still in the stage of descriptive analysis, and in the following 20 years, there has been no significant progress in the study of financial crisis prediction. After 1990, related research began to change from simply studying corporate financial variables to predict corporate financial risks, to adding variables such as macroeconomic environment to the study for further research. In the subsequent series of studies, some scholars such as Odom and Coats tried to use the artificial neural network (ANN) to conduct research. From the obtained results, the success rate is relatively low. Only Coats (1993) used Altman to discriminate on the basis of the five judgment variables proposed by the Z-Score model and obtained relatively good judgment results in the prediction of crisis enterprises. The accuracy rate is 91%, and the accuracy rate for normal enterprise prediction is 96%, which is slightly better than the discriminant analysis model. So far, the research on financial early warning in China is not very sufficient, and it is basically in its infancy [14–17]. Among the scholars who studied, most of them introduced foreign theories in terms of research in their own countries. Some of them discussed theories and feasibility and put forward suggestions and measures in system construction and legislation, and very few conducted in-depth research on issues such as models.

This paper selects the financial indicator data of 60 companies from 2017 to 2019 and divides it into training samples and prediction samples for empirical research. We use the SPSS20.0 software to screen out model indicators for the next step of model building, optimize variables by factor analysis, preprocess the data, and establish and compare Logistic models using financial indicators and adding nonfinancial indicators. Matlab software is used to train the training samples to establish a BP neural network early warning model, and the prediction samples are used to test the accuracy of the model prediction [18–20].

2. Methods and Theory

2.1. Design Steps. The structure of the financial early warning model can be divided into five steps as shown in Figure 1. First, we determine the research sample type. The financial early warning has certain limitations. Different periods and different industries will correspond to different financial characteristics. Determining the sample type and enterprise characteristics is research. The second step is to select index variables according to the characteristics of different industries and adjust them appropriately. Third, we need to screen the financial indicators and select the indicators required to build the model. The screening methods in

existing research generally include factor analysis and principal component analysis method. Fourth, we establish the model. This paper uses Logistic regression and BP neural network model to establish financial early warning models, respectively, to evaluate the warning level.

2.2. Logistic Early Warning Model Design

2.2.1. Research Sample. In this paper, 60 companies from 2017 to 2019 are selected first, and the sample companies are screened and classified according to the classification standard of early warning degree. When classifying the sample data, we try to maintain the balance between the samples. The specific classification is shown in Table 1.

2.2.2. Early Warning Indicator Variables. Preliminary selection is made from two aspects: financial indicators and nonfinancial indicators. Financial indicators are selected from five aspects: profitability and income quality, solvency and capital structure, operating capacity and cash flow, development capacity, and risk level. Nonfinancial indicators include related party transactions, violations, audit reports, litigation, and arbitration. We select indicators that can truly reflect the financial risks of the industry and then conduct further screening. Through the normality test of the indicators and the significance test, the indicators selected in this paper are shown in Table 2.

Factor analysis requires a high degree of correlation between variables, and a correlation test needs to be carried out before factor analysis. Financial indicators are calculated through financial data, and there is a high degree of correlation between variables. Nonfinancial indicators do not need factor analysis, so we incorporate it directly into the variables of the final model building.

Through the factor analysis method of the KMO and Bartlett sphere test, 11 financial indicators are preliminarily screened to ensure that the screened financial variables are suitable for the empirical requirements of companies. The data are obtained through SPSS analysis software to obtain the 11 indicators for factor analysis. Through factor analysis, it is concluded that the eigenvalues of the first four common factors exceed 1, so the first four common factors can be used as surrogate variables for 11 financial indicators. The public factor F1 can be explained by variables 3, 4, 5, and 6, reflecting the solvency and capital structure. The public factor F2 can be explained by 1, 2, 9, and 10, reflecting the profitability and growth ability. The public factor F3 can be explained by variables 7 and 8, reflecting the level of operating capacity and cash flow. The public factor F4 can be explained by variable 11, reflecting the risk level of the enterprise.

2.2.3. BP Neural Network Early Warning Model Design. The realization of the BP neural network financial crisis early warning model is divided into input and output layer variable design, hidden layer node design, error and learning rate selection, etc. The realization process is as follows:



FIGURE 1: Design steps of financial early warning model.

TABLE 1: Sample data grouping.

Warning interval	Training data	Prediction data	Total
Safety warning zone	40	20	60
Light warning range	20	10	30
Severe warning range	10	5	15
Total	70	35	105

TABLE 2: Financial indicator variables that passed the test.

Number of variables	Indicator name
1	Total operating cost ratio
2	ROA
3	Assets and liabilities
4	Capital adequacy ratio
5	Current ratio
6	Quick ratio
7	Current asset turnover
8	Accounts Receivable Turnover
9	Asset cash ratio
10	Net asset growth rate
11	Financial leverage
12	Audit report

Note. Variable 12 is a nonfinancial indicator, the rest are financial indicators.

- (1) *Select Training and Prediction Samples.* According to the research direction, the financial samples of enterprises are selected, the two-year financial data of 40 companies are used as training samples for training, the data of 20 companies are used for prediction, and the predicted value of the model is compared with the expected output value.
- (2) *Determine the Input and Output Layer Changes.* In general, the number of input nodes is positively correlated with the approximation result of the neural network. The more input neuron nodes, the better the approximation effect. However, when irrelevant indicators are incorporated into the neural network training, the fitting will occur due to the increase in the number of indicators. Too many nodes will increase the calculation amount of the BP neural network, which will lead to longer model training time, affecting the output results.

According to the indicator variables selected by the Logistic early warning model, 11 financial indicators and 1 nonfinancial indicator are used as the input vectors of the model and the number of input variables is confirmed; that is, the input layer of the model is positioned as 12 in this paper.

The output layer uses the expected output value after training and the classification result of the

financial warning degree of energy enterprises as the output vector for training. Some studies have concluded that the number of nodes in the prediction output layer of the financial crisis is set to 1.

- (3) *Determine the Number of Hidden Layer Nodes.* The model is more complicated to determine the number of nodes in the hidden layer, which is generally set by empirical values. According to the comprehensive consideration of the sample size and the number of variables, setting the appropriate number of nodes is directly related to the results of the model. Too many hidden layer nodes will increase the learning practice, and too few will reduce the collection ability, and it is impossible to find the inherent laws of the data through self-learning.

The number of hidden layer nodes is different for different models, and the following are selected from common empirical formulas, where n is the number of input nodes and m is the number of output nodes:

$$L = \log_2 n, \quad (1)$$

$$L = \frac{(m \times n)}{2},$$

$$L = \sqrt{m + n} + a, \quad (2)$$

$$L = n + 0.618(n - m). \quad (3)$$

In this paper, a is a constant between 1 and 10. Empirical formula (2) is selected. First, the input node $n = 12$ is brought into the model, and a tentative test is carried out, and it is continuously adjusted according to the results. The final number of nodes is selected as 7.

3. Results and Discussion

3.1. Establishment and Prediction of the Logistic Model of Financial Indicators

3.1.1. Establishing a Logistic Model of Financial Indicators. We add nonfinancial indicator variables and four common factors for the likelihood ratio test, and the specific values are shown in Table 3.

The comprehensive indicator early warning model refers to adding the nonfinancial indicator variable 12 to the pure financial indicators. It is further concluded that the Sig. values are less than 0.05 and pass the significance test. The fitting effect of this model is good.

3.1.2. Comprehensive Index Logistic Model Fitting Test. We use the corporate financial data as test data to test the comprehensive index Logistic model. The results are shown

TABLE 3: Likelihood ratio test table.

Effect	Standard			Likelihood ratio test		
	Simplified model AIC	Simplified model BIC	$-2x$ log-likelihood of the reduced model	K^2	df	Significant level
Intercept	100.34	129.41	74.32	0	2	0
FAC1_1	129.84	155.67	102.41	24.32	2	0
FAC2_1	174.69	200.31	161.53	74.37	2	0
FAC3_1	70.13	97.95	54.31		2	0
FAC4_1	78.85	100.24	58.51		2	0
12	75.35	97.65	52.46		2	0

in Table 4. After adding nonfinancial indicators, the obtained comprehensive Logistic model is fitted and tested, and the test results are safe, light, and severe. The prediction accuracy is 90.00%, 80.00%, and 90.00%, respectively, and the overall fitting test result was 87.00%.

Using financial data as forecast data to test the comprehensive index Logistic model, the results are shown in Table 5. The prediction accuracy of the Logistic financial early warning model after adding nonfinancial indicators for safety, light, and severe is 95%, 60.0%, and 80.0%, respectively, and the overall prediction test result is 78.0%.

3.2. Model Training and Prediction. After the design and adjustment of the BP neural network model, the training of the sample data is started, so that the network can distinguish the output vector of the financial warning area of the enterprise. In this paper, combined with the number of nodes in the hidden layer above, according to the formula, multiple solutions are performed to obtain multiple values, and the bring-in test is carried out in turn. According to the size of the error, when the hidden layer is 7, the MSE minimum fitting effect is the best. When the number of nodes with layers is set to 7, the sample training error curve is shown in Figure 2.

The training is carried out according to the results of the early warning area of the training sample. According to the comparison between the training value and the actual value in Figure 3, the error between the fitting value and the actual value is not large, and the fitting degree is good. The prediction accuracy of the training value is shown in Table 6.

The predicted test results obtained from the predicted samples are shown in Figure 4.

According to the results of training and equations, it can be judged that when the hidden layer is 7, the fitting effect and prediction effect of the model established in this paper are the best, and the overall test results are 100% and 93%, respectively. The fitting result of the BP neural network can be calculated to be optimal, and the fitting result for the three-year financial warning result of the enterprise reaches 100%. It can be said that when the hidden layer node is set to Table 7, the effect of the model is optimal. The prediction test is then carried out on the financial data of the prediction sample. According to Table 7, it can be seen that the prediction effect is slightly worse than the fitting result. For the mild early warning interval, the overall prediction test result is 93%.

TABLE 4: Comprehensive index Logistic model fitting test results.

Observed value	Predicted value			Percent correction (%)
	Safe	Light	Severe	
Safety warning zone	36	4	0	90.00
Light warning range	3	16	1	80.00
Severe warning range	0	1	9	90.00
Total percentage	56%	30%	14%	87.00

TABLE 5: Prediction test results of the Logistic model for comprehensive indicators.

Observed value	Predicted value			Percent correction (%)
	Safe	Light	Severe	
Safety warning zone	19	1	0	95.00
Light warning range	2	6	2	60.00
Severe warning range	0	1	4	80.00
Total percentage	60%	23%	17%	78.00

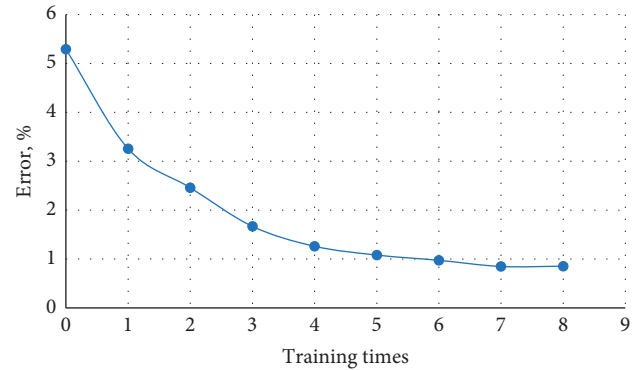


FIGURE 2: Sample training error curve.

When the number of hidden layer nodes is small, the acquisition ability is poor, the learning speed is slow, and the accuracy rate is relatively low. When the number of nodes increases, the accuracy rate is improved. Therefore, the parameters of the model selected in this paper are detailed in Table 8.

3.3. Comparative Analysis of Financial Early Warning Models. Financial early warning research is mainly divided into qualitative and quantitative research. Qualitative research

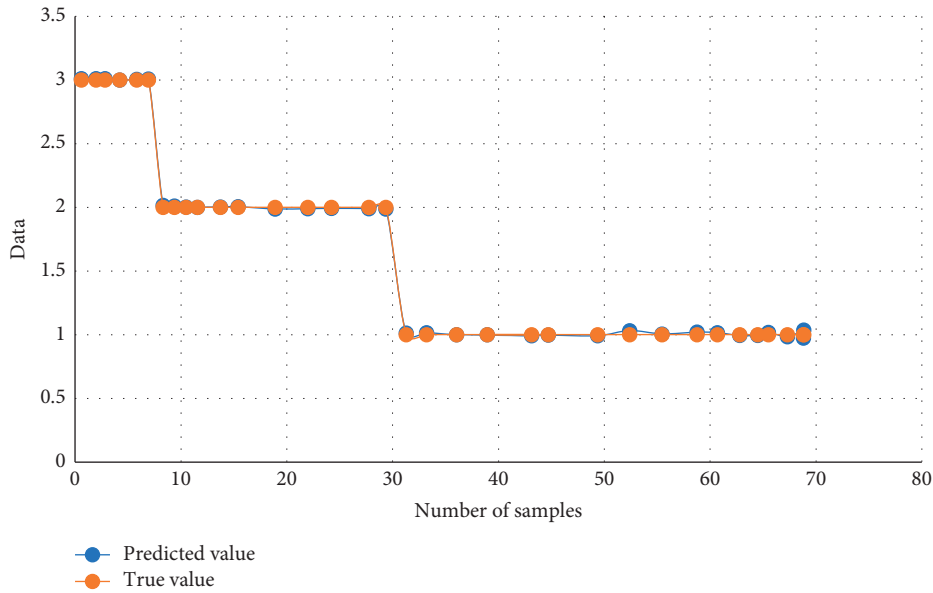


FIGURE 3: Comparison between training and actual data.

TABLE 6: Classification diagram of fitting results.

Observed value	Predicted value			Percent correction (%)
	Safe	Light	Severe	
Safety warning zone	40	0	0	100.0
Light warning range	0	20	0	100.0
Severe warning range	0	0	10	100.0
Total percentage	57.1%	28.6%	14.3%	100.0

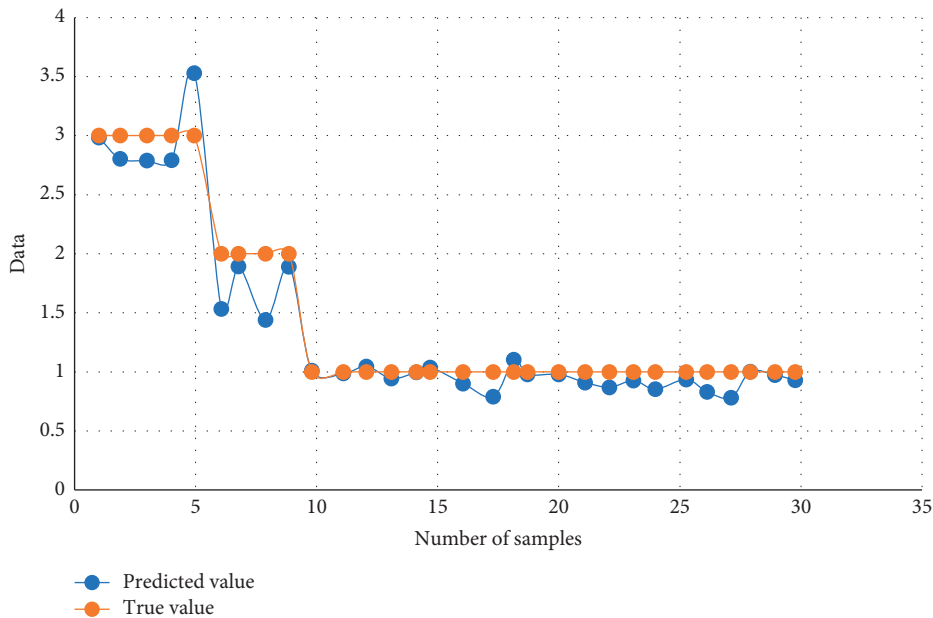


FIGURE 4: Comparison between training and actual data.

often uses financial professionals to make subjective judgments, but the results are subjective and lack objectivity and fairness. Quantitative research mainly analyzes financial

data, reduces the subjectivity of human judgment, and improves the accuracy of prediction through mathematical models.

TABLE 7: Classification of prediction test results.

Observed value	Predicted value			Percent correction (%)
	Safe	Light	Severe	
Safety warning zone	20	0	0	100.00
Light warning range	2	8	0	80.00
Severe warning range	0	0	5	100.00
Total percentage	63%	23%	14%	93

TABLE 8: Parameters of the BP neural network financial early warning model.

Model design	Parameter
Network level	Number of nodes
Input layer	12
Hidden layer	7
Input layer	1
Transfer function	Sigmoid
Training function	Trainlm
Learning rate	0.01
System error	0.001
Number of iterations	100,000

In practice, financial early warning usually adopts a combination of qualitative and quantitative research methods. Through the methods established by various models, the early warning results are closer to reality and more accurate. Traditional mathematical statistics need to set assumptions and constraints, which cannot fit the reality well, and artificial neural network can better make up for this problem. The comprehensive logistic model with qualitative indicators is not as accurate as the BP neural network. Simple mathematical statistical models have higher requirements on data and more restrictions on the use of models. The correlation between indicators needs to be considered when building models. BP neural network has no constraints on data indicators and has self-learning and feedback capabilities. These advantages have greatly improved the accuracy of financial early warning models.

3.4. Establishing a Financial Early Warning System. In recent years, with the continuous changes in the industry, many investors have entered the Chinese market, and the formation of a diversified pattern has also brought huge challenges to Chinese enterprises. The enterprise system is huge and the organizational structure is complex. It is difficult to avoid risks only by relying on the financial early warning model. It is imperative to establish a corresponding early warning system. Chinese enterprises pay more and more attention to the use of scientific methods for management, and managers need to grasp the operating conditions and financial conditions. At this time, the financial early warning has become one of the important management methods to monitor enterprise risks. The financial early warning system can be applied to the business activities of various capital transactions of the enterprise. It is a kind of research based on various financial-related data of the

enterprise over the years or peers and makes full use of the theoretical foundations of finance, management, etc. When the enterprise falls into a financial crisis, an early warning signal is issued in time to notify managers to be more vigilant. An excellent financial early warning system suitable for enterprises should include the collection and processing of financial information, supervision, and prevention analysis of financial risks. The collection and processing of data and information is a prerequisite for establishing an early warning system, supervision is the core part, and analysis and response are the keys to risk early warning.

4. Conclusion

- (1) This paper collects the financial indicators of enterprises from 2017 to 2019, and the data mainly come from the stock exchanges and the CCER database. Three criteria are selected as the basis for dividing the early warning degree of energy companies, and the financial data of 40 companies in the same accounting period of 2 years are selected for research, and then the data of 20 companies are selected as prediction test samples for testing.
- (2) Comparing the Logistic model with the BP neural network, the BP neural network does not have strict requirements on financial indicators in terms of constraints. According to the training results and prediction results, the BP neural network is better than the Logistic model. The calculation of the Logistic model is relatively complex, but it can be expressed in mathematical expressions. BP neural network has strong self-learning ability and computing ability, but it cannot write specific equations. According to the conclusion of the municipal research carried out in this paper, BP neural networks are more suitable for financial early warning of Chinese enterprises.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The authors would like to show sincere thanks to those technicians who have contributed to this research.

References

- [1] M. Bussiere and M. Fratzscher, "Towards a new early warning system of financial crises," *Journal of International Money and Finance*, vol. 25, no. 6, pp. 953–973, 2006.
- [2] H. J. Edison, "Do indicators of financial crises work? an evaluation of an early warning system," *International Journal of Finance & Economics*, vol. 8, no. 1, pp. 11–53, 2003.

- [3] K. J. Oh, T. Y. Kim, and C. Kim, "An early warning system for detection of financial crisis using financial market volatility," *Expert Systems*, vol. 23, no. 2, pp. 83–98, 2006.
- [4] C. Diks, C. Hommes, and J. Wang, "Critical slowing down as an early warning signal for financial crises?" *Empirical Economics*, vol. 57, no. 4, pp. 1201–1228, 2019.
- [5] D. H. Kim, S. J. Lee, K. J. Oh, and T. Y. Kim, "An early warning system for financial crisis using a stock market instability index," *Expert Systems*, vol. 26, no. 3, pp. 260–273, 2009.
- [6] H. Dastkhan, "Network-based early warning system to predict financial crisis," *International Journal of Finance & Economics*, vol. 26, no. 1, pp. 594–616, 2021.
- [7] M. Jemović and S. Marinković, "Determinants of financial crises—an early warning system based on panel logit regression," *International Journal of Finance & Economics*, vol. 26, no. 1, pp. 103–117, 2021.
- [8] L. Wu, L. Chen, and X. Hao, "Multi-sensor data fusion algorithm for indoor fire early warning based on BP neural network," *Information*, vol. 12, no. 2, p. 59, 2021.
- [9] G. Li, "An integrated model of rough set and radial basis function neural network for early warning of enterprise human resource crisis," *International Journal of Fuzzy Systems*, vol. 21, no. 8, pp. 2462–2471, 2019.
- [10] Y. Cao, X. Chen, D. D. Wu, and M. Mo, "Early warning of enterprise decline in a life cycle using neural networks and rough set theory," *Expert Systems with Applications*, vol. 38, no. 6, pp. 6424–6429, 2011.
- [11] B. Yang, L. X. Li, H. Ji, and J. Xu, "An early warning system for loan risk assessment using artificial neural networks," *Knowledge-Based Systems*, vol. 14, no. 5-6, pp. 303–306, 2001.
- [12] J. Chen, Y. Xu, S. Xu, C. Zhao, and H. Chen, "Research on construction of anti-dumping early warning model based on BP neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5649–5659, 2020.
- [13] J. Paulsen, "Risk theory in a stochastic economic environment," *Stochastic Processes and their Applications*, vol. 46, no. 2, pp. 327–361, 1993.
- [14] J. L. Tobin and S. Rose-Ackerman, "When BITs have some bite: the political-economic environment for bilateral investment treaties," *The Review of International Organizations*, vol. 6, no. 1, pp. 1–32, 2011.
- [15] C. Chidoko, G. Makuyana, P. Matungamire, and J. Bemani, "Impact of the informal sector on the current Zimbabwean economic environment," *International Journal of Economic Research*, vol. 2, no. 6, pp. 26–28, 2011.
- [16] V. Svata, "IS audit considerations in respect of current economic environment," *Journal of Systems Integration*, vol. 2, no. 1, pp. 12–20, 2011.
- [17] D. V. Kamensky, O. O. Dudorov, R. O. Movchan, A. A. Vozniuk, and T. P. Makarenko, "Insider trading in the global economic environment: elements of criminal liability," *International Journal of Management*, vol. 11, no. 12, pp. 1679–1688, 2020.
- [18] S. Manca and M. Delfino, "Adapting educational practices in emergency remote education: continuity and change from a student perspective," *British Journal of Educational Technology*, vol. 52, no. 4, pp. 1394–1413, 2021.
- [19] S. Ding, C. Su, and J. Yu, "An optimizing BP neural network algorithm based on genetic algorithm," *Artificial Intelligence Review*, vol. 36, no. 2, pp. 153–162, 2011.
- [20] L. Huipeng, "Martial arts competitive decision-making algorithm based on improved BP neural network," *Journal of Healthcare Engineering*, vol. 2021, no. 8, Article ID 9920186, 8 pages, 2021.

Research Article

Research on Intelligent Scheduling Scheme of Aerobics Competition for Multi-Intelligent Decision-Making

Hongjing Guan ¹, Zhi Tan,² and Xingrui Zhao¹

¹Graduate School, José Rizal University, Manila 0900, Philippines

²Graduate School, Woosuk University, Jeonju 514500, Republic of Korea

Correspondence should be addressed to Hongjing Guan; hongjing.guan@my.jru.edu

Received 27 February 2022; Revised 10 April 2022; Accepted 19 April 2022; Published 23 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Hongjing Guan 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.

Multi-intelligent decision-making has a good development at present. Based on a series of technologies such as artificial intelligence, multi-intelligent decision-making is involved in many aspects, and the country also attaches great importance to the development of such science and technology. At the same time, at present, the development of the sports industry is not balanced. Even if the country values the sports industry, it needs investment in science and technology. This paper studies the scheduling scheme of aerobics competition in sports competition. By introducing the design scheme of a multi-intelligent decision-making system and improving the MFDRL-CTDE algorithm, the similarity between the action sequence of participants in aerobics competition and the standard action sequence is obtained. Three algorithms including Markov, MFDRL-CTDE algorithm, and improved MFDRL-CTDE algorithm are used for simulation experiments, and the improved MFDRL-CTDE algorithm is more effective and stable for the decision-making of aerobics competition.

1. Introduction

Our country is very interested in the development of science and technology. The application of intelligent decision-making system has penetrated into all walks of life, and the application of intelligent decision-making system has also been internationally recognized. Researchers have developed related intelligent decision-making in depth, which greatly improves the performance of various related decision-making algorithms and makes related applications more efficient. While taking care of the development of science and technology, the country pays more attention to the development of sports, especially the development of aerobics competition. The development history of aerobics is not long ago, but it is loved by the broad masses of people, and it has the performance of sports competition while having an aesthetic feeling. However, the development is not balanced, so it is necessary to put science and technology into aerobics competition and intelligent decision-

making into aerobics competition, so as to make aerobics competition have better development.

In the process of studying intelligent decision support system, many new methods have emerged [1]. According to the structural characteristics and system characteristics of the system, an improved model can be designed well. Group is the foundation of the organization of individual intelligence, and the model and strategy discussed are also based on this characteristic. Combined with the combination mechanism, the model can run efficiently. The decision-making process plays an important role in the interactive trend of multi-intelligence combination [2]. In order to accurately predict, it is impossible to use probability or theory alone. What game can bring is very good decision-making and identification. The Bayesian game model is the combination of game and graph, which can solve the complex situation caused by multi-intelligent agent interaction. Compared with other intelligent decision information systems, this model has better decision efficiency. Because decision-making and evaluation are not carried out

at the same time, their operation process is affected by many factors. Nevertheless [3], it is extremely urgent to improve the efficiency of group decision-making. Therefore, the rational political model is put forward, and its purpose is to improve the operational efficiency of group decision-making in the face of uncertain factors. In this process, this model clarifies the factors. Moreover, this model obtains the best scheme of group decision-making on the basis of a sequence framework when there are different complete evaluations, information, and opinions. In real-time work [4], such as the intensive care unit, what doctors need is the result of matching time and information, and the important value of an intelligent decision support system is the agent of two or more people to complete tasks cooperatively. At present, the architecture of a multiagent system has been proposed, which is undoubtedly to support the decision-making of clinical data and predictive status required by doctors. The artificial neural network algorithm is divided into two steps [5]. The first step is to use the improved algorithm to calculate the relevant weight matrix. The second step is to apply the method to real system decision-making, such as a multi-intelligent robot soccer system. The final simulation results show that this technology can be efficiently used in small- and medium-sized league system decision-making. Multi-criteria decision-making method can be applied to aerospace systems effectively [6] and help decision makers to solve related conflict problems. In the face of a variety of multicriteria decision-making methods, we need to choose the appropriate method, but this is a very complex problem related to multicriteria decision-making. The proposal of 14 criteria can evaluate the feasibility of the multicriteria decision-making method, so that the preference of decision makers can be better informationized. If we want to develop an intelligent knowledge system, we can optimize the suitability index. In order to prove the effectiveness of the system, the aircraft selection problem can solve this problem well. In order to realize the UniComBOS of intelligent decision support system [7], we propose to use verbal decision to analyze the paradigm. In this range, for multi-standard alternatives, the processing ability of individual decision makers is considered, and the information of decision makers' psychological preferences is accurately extracted on this premise. In the user interface, the decision-maker's psychological preference is distinguished in the form of graphic color, which provides a psychological comparison for the decision-maker and can test whether the decision-maker's answer is consistent. Like conventional decision-making tools, this scheme decomposes the whole part with the least number of criteria, so that the comparison range of decision makers can be reduced to a single criterion, and then, the criteria are added until the advantages of the optimal substitute are found. However, in the case of multiple standards, a single optimal alternative cannot be found by the system, so only a group of alternatives can be found. However, there is no comparability between these alternatives, but they still have great advantages over those that are not selected. In the EFQM excellence model, literature [8] is used to achieve business excellence, which is mostly used for the self-business evaluation of many

nonlarge enterprises in Britain. In fact, this is a problem related to multistandard decision analysis. At present, experiments have proved that in order to realize the method of achieving excellent business through action, an intelligent decision system is used to support program groups, which not only makes the average score come out but also can wait for numerical results and graphic comparison. At present, the society is interested in sports as well as the development of science and technology [9]. In Korea, aerobic exercise is popular with people. However, it has been emphasized recently that sports need technology and competition, which makes the development trend of aerobic exercise that shows an unbalanced trend. Among them, the development of aerobics is relatively rapid. In order to find a new way out for aerobics, it is very important to understand its history and introduction process. The demand and rapid change of society for aerobics is the key to the development of aerobics, and the development needs to be carried out in a moderate way and cannot be too radical in any field. According to the data shown, the relevant results of aerobics competition need to be analyzed with various related techniques [10], which can well analyze whether the scores given by the referee are reasonable and consistent. It makes the selection and scoring of referees more scientific and effective. Referees need to maintain an objective and fair attitude towards competition scoring, and rational evaluation can better develop and build the future of aerobics competition. Aerobics Championship needs to analyze the movements [11]. According to the rules, difficulties, and characteristics of the world aerobics competition, each group performs differently. There are outstanding performances in dynamic strength or outstanding performances in balance flexibility, and the most outstanding ones in each group are different. For the analysis of aerobics competition results, we can use related research methods, such as comprehensive omission research and variance analysis [12]. Through the experimental data, we can observe that the scores given by the judges are relatively the same, and of course, the performance displayed by the finished judges is better than that of the artistic judges. Therefore, no matter what kind of aerobics competition, it needs a systematic scoring test program. Relevant technologies are used to simplify scores and evaluate objectively. At the same time, participating members can question scores and arbitrate. According to the statistical analysis of the skills of the first six participants in a certain aerobics competition [13], and summarizing the techniques of music rhythm and aerobic exercise, we can develop ladder aerobics and find out its characteristics. The most important point is that this can make junior coaches have more references for aerobics competition. For students' aerobics competition, on-site observation, interview, and video analysis of the later competition are also very important [14]. This is a strengthening of aerobics education for students and a reference for the competition itself. Our country attaches great importance to the development of gymnastics and aerobics [15], and many related training competitions are held by national-level organizations. However, the future development of aerobics needs to be

analyzed through data literature, refers to recent training, and changes to solve the problems on the road of development.

2. Multi-Intelligent Decision Analysis

2.1. Basic Concepts. Multi-intelligent decision-making refers to the collective integration of multiple intelligent units, so a huge system is composed of multiagents. One of the main tasks of multi-intelligent decision-making is that decision makers disassemble comparatively complex systems, which is a decision-making process. A complex system is divided into several small systems, and these small systems will be combined again. It should be noted that these small systems are interactive, which is why we can ensure that each system is interactive, so as to improve work efficiency. Compared with other technologies, multi-intelligent decision-making has absolute advantages, which are manifested in its independence, relevance, and application development. The classification of intelligent decision is shown in Figure 1.

2.2. Markov. The sequential decision means that after the t -th iteration, the state s of the environment will be received by the agent, and on this basis, an action a will occur. Because the agent and the environment are related to each other, this means that the previous actions will also affect the environment, which will make the agent get a reward, that is, a reward R_{t+1} , so as to stimulate its new state. It is precisely because of the interaction between agent and environment that sequences can be generated. Markov's decision process is to formulate the above process, and Markov's definition formula is as follows:

$$P[S_{t+1}|S_t] = P[S_{t+1}|S_1, \dots, S_t]. \quad (1)$$

That is, in Markov, the state of the $t + 1$ iteration and the state of the t iteration are related.

State s and the state transition probability p based on its next state s' are expressed as follows:

$$P_{ss'} = P[S_{t+1} = s' | S_t = s]. \quad (2)$$

That is, in Markov, the state of the $t + 1$ iteration is only related to the state of the t iteration.

The formula for getting the reward expectation when the state s moves to the next state s' based on it is as follows:

$$R_s = E\left[\frac{R_{t+1}}{S_t = s}\right]. \quad (3)$$

The formula for all rewards and G from state s to the final state point is as follows:

$$G_t = R_{t+1} + \gamma R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k} + k + 1, \quad (4)$$

where γ is a discount factor, which is a parameter used to express attenuation.

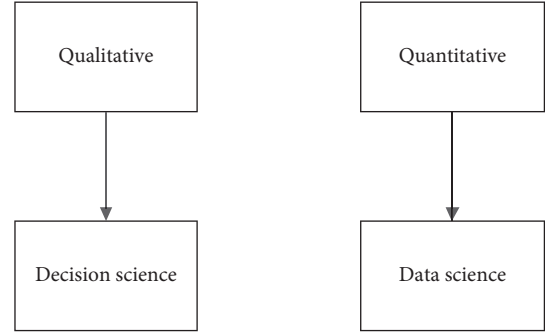


FIGURE 1: Classification diagram of intelligent decision.

2.3. MFDRL-CTDE Algorithm

2.3.1. Multiagent Fuzzy Deep Reinforcement Learning. The main task of a fuzzy inference system is to fuzzify and defuzzify variables. Its core idea is that when the state is input, variables are continuously fuzzified, and the fuzzification of behavior is due to the prescribed rules, and finally, the behavior should be defuzzified and output. When the output variable is controlled at $[0, 1]$, the input fusion weight w_F of the system is the combination of the cumulative average reward and the sample priority in the iteration period, and its normalization formula is as follows:

$$\bar{w}_F = \frac{w_F}{\sum_i w_{F_i}} = [\bar{w}_{F_1}, \bar{w}_{F_2}, \dots, \bar{w}_{F_i}, \dots, \bar{w}_{F_I}], \quad (5)$$

where I is the total number of agents.

However, the network parameters θ in the system will not be equal to 0 after many iterations, so the formula for updating the parameters is as follows:

$$\begin{aligned} \theta_{V \text{ tar}} &= \sum_i \bar{w}_{F_i} \theta_{V \text{ val}}^i, \\ \theta_{A \text{ tar}} &= \sum_i \bar{w}_{F_i} \theta_{A \text{ val}}^i, \\ \theta_{\text{tar}} &= \sum_i \bar{w}_{F_i} \theta_{\text{val}}^i. \end{aligned} \quad (6)$$

2.3.2. Centralized Training and Decentralized Execution. Each agent affects the value function that determines each of them, which is indispensable in multi-intelligence deep reinforcement learning. Because of this, Markov's decision process in multi-intelligence system is not working well, and the hidden information is difficult to be fully discovered by the agent, so it is necessary to capture this information during training. Centralized training and decentralized execution make use of this point to help agent improve efficiency and, at the same time, store the available information of agent interaction into the shared experience pool, thus further improving efficiency.

For centralized training, the value functions of agents are related to joint behaviors rather than local ones, and they are

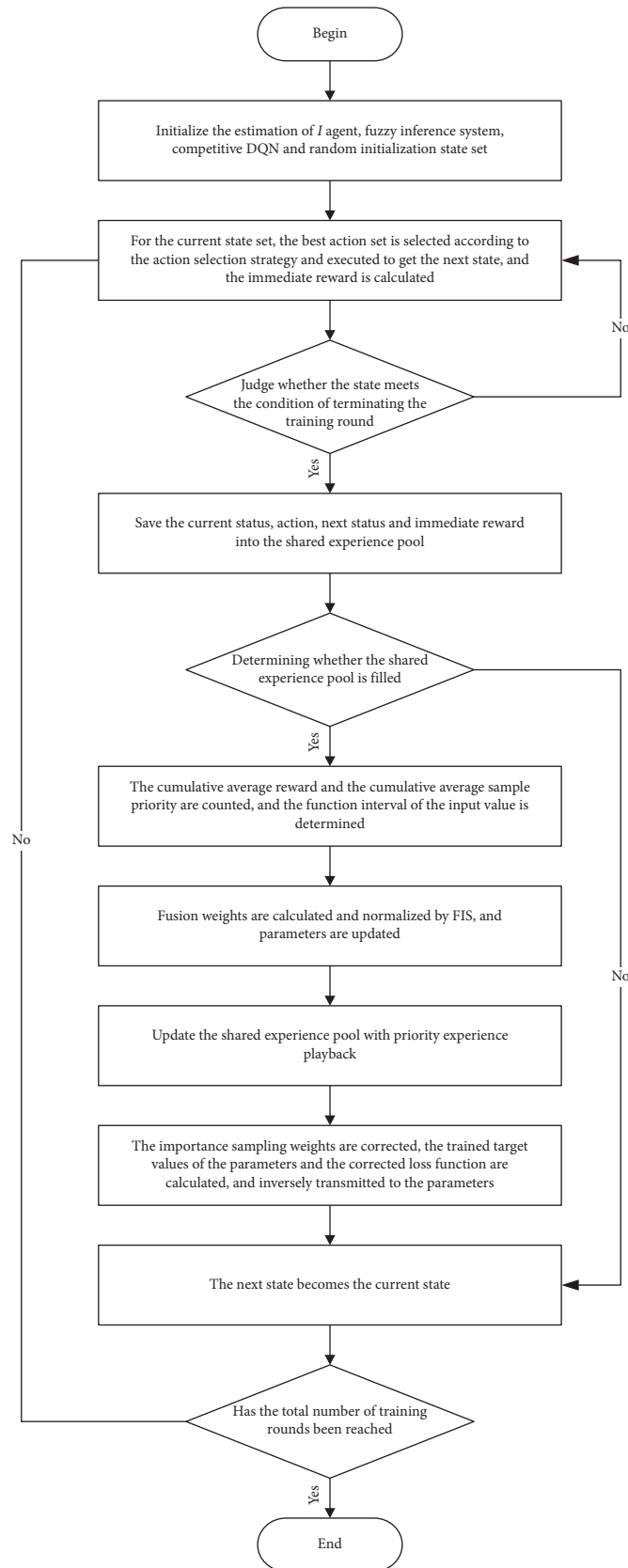


FIGURE 2: Flow chart of MFDRL-CTDE algorithm.

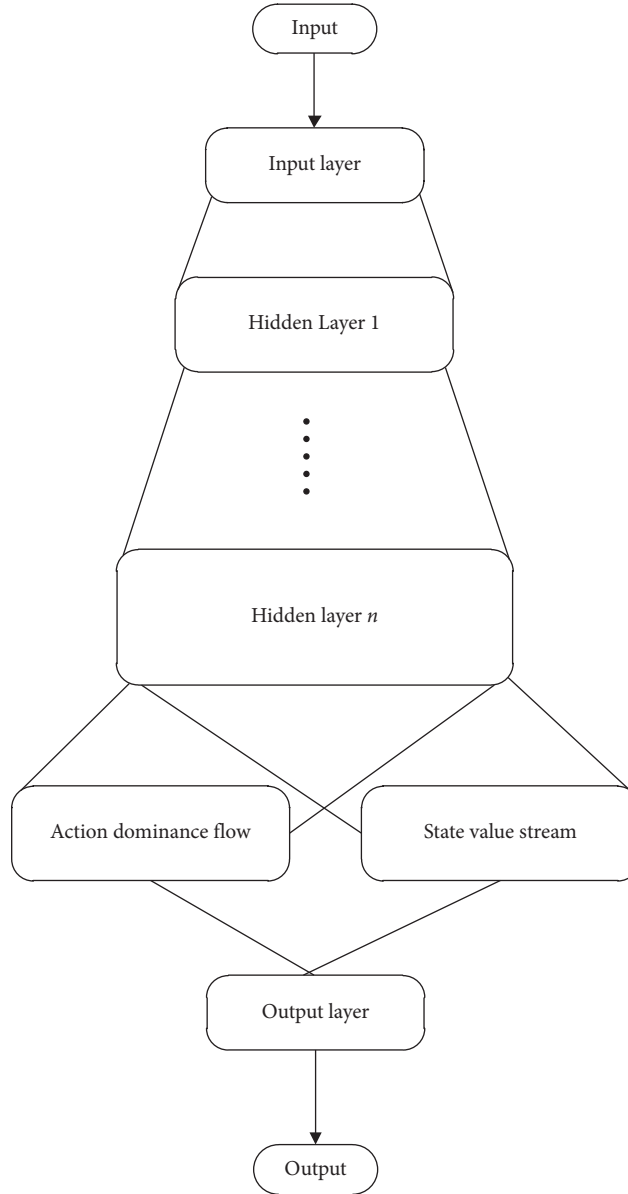


FIGURE 3: Competitive DQN structure diagram.

trained with extra information. For decentralized execution, agent decision-making only depends on the information that has been mined, rather than all the complete information.

3. Optimization Algorithm

3.1. Optimize MFDRL-CTDE Algorithm. Because of the low fitting ability of the value function in the initial stage, the MFDRL-CTde algorithm cannot guarantee its high efficiency. Therefore, in order to improve the efficiency of the algorithm and reduce the volatility of the algorithm, it is necessary to add competitive DQN and preferential empirical playback. The algorithm flow is shown in Figure 2.

3.1.1. Action Selection Strategy. In order to solve the problem that the agent may choose to execute random actions caused by the constant or decreasing parameter ϵ of a typical action selection strategy in the later stage, thus greatly reducing the convergence speed of the algorithm, an optimized action selection strategy is needed, and the formula is as follows:

$$\pi^i(s_t^i) = \begin{cases} a_{\text{random}}, & \text{rand} < \epsilon, \\ \max_{a'} Q_t^i(s_t^i, a'), & \text{rand} \geq \epsilon, \end{cases} \quad \epsilon = \frac{1}{\lambda \bar{E}/T} \quad (7)$$

where $\pi^i(s_t^i)$ is the best action strategy of the agent belonging to the i -th in-state s_t^i when the t -th iterative update is carried

TABLE 1: Test function calculation results table.

Algorithm		f_1	f_2	f_3	f_4	f_5	f_6
Markov	Average	1.17E + 03	1.6832	7.4741	4.80E + 02	0.0357	2.4201
	Standard deviation	1.99E + 03	3.8855	6.2378	2.08E + 02	0.0543	1.2608
MFDRL-CTDE	Average	1.58E + 02	1.6203	8.3282	3.32E + 01	0.0243	2.4001
	Standard deviation	1.70E + 03	3.4943	5.2941	1.8931	0.0083	1.0031
Improved MFDRL-CTDE	Average	1.32E + 02	1.5983	6.2984	3.43E + 01	0.0123	2.1031
	Standard deviation	1.76E + 02	3.0523	5.3823	0.0342	0.0031	0.3134

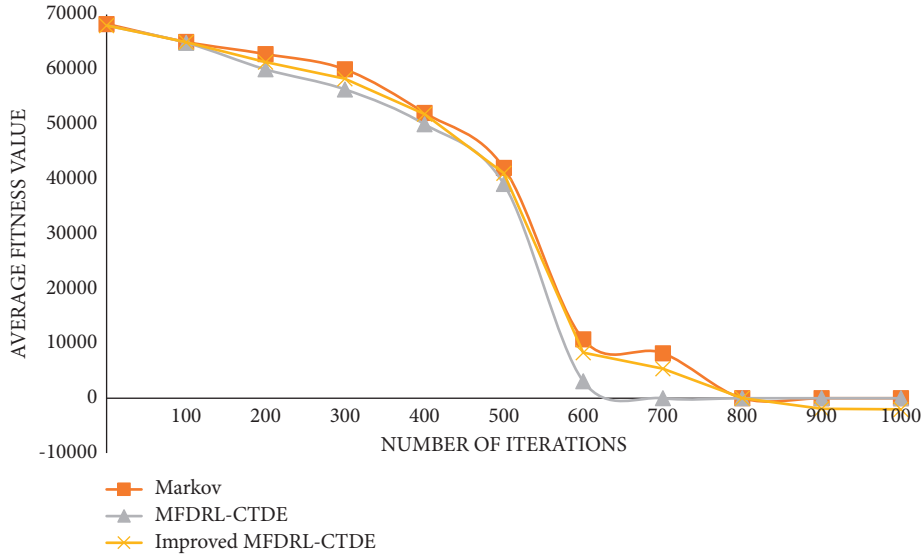


FIGURE 4: Graph of function 1.

out, a_{random} is the random selection action, rand is the random number with a value range of $[0, 1]$, λ is the control rate of ε descent, \bar{t} is the current number of training rounds, and \bar{T} is the total number of training rounds.

3.1.2. Competitive DQN. In deep reinforcement learning, DQN is undoubtedly the most frequently used method, but it is only applicable to a single agent. Facing multiagent, it is obvious that the efficiency of ordinary DQN cannot keep up with a large number of state and action combinations of each agent. Therefore, for multi-intelligent systems, it is necessary to apply competitive DQN as the basic structure to improve the original structure and improve the algorithm efficiency.

Competitive DQN solves the problem that a certain state is immune to any action, which can not affect the subsequent state. It breaks the value combination of state and action, which can improve efficiency. The competitive DQN structure is shown in Figure 3.

Among them, the state value function will have a value equal to 0, and the action advantage function will also have a value equal to 0. In order to avoid this kind of situation, it is necessary to reduce the Q value, but the order of the action dominance function cannot be changed, so as to reach the degree of freedom without redundancy. The formula is as follows:

$$Q(s_t, a_t; \theta, \theta_V, \theta_A) = V(s_t; \theta, \theta_V) + \left[A(s_t, a_t; \theta, \theta_A) - \frac{1}{N_A} \sum_{a'} A(s_t, a'; \theta, \theta_A) \right], \quad (8)$$

where $Q(s_t, a_t; \theta, \theta_V, \theta_A)$ is the agent in state s_t that executes the Q value of action a_t at the t -th iterative update. $V(s_t; \theta, \theta_V)$ is the state value function, that is, the value of the state itself, $A(s_t, a_t; \theta, \theta_A)$ is the action dominance function, that is, the extra value generated after the corresponding action is selected, and N_A is the number of actions that are likely to occur.

3.1.3. Priority Experience Is Put Back. In the ordinary MFDRL-CTDE algorithm, random sampling is adopted for the step of putting back experience, which undoubtedly subtracts the playback efficiency. Therefore, in order to improve the efficiency of using samples and enhance the function of experience playback, it is necessary to introduce

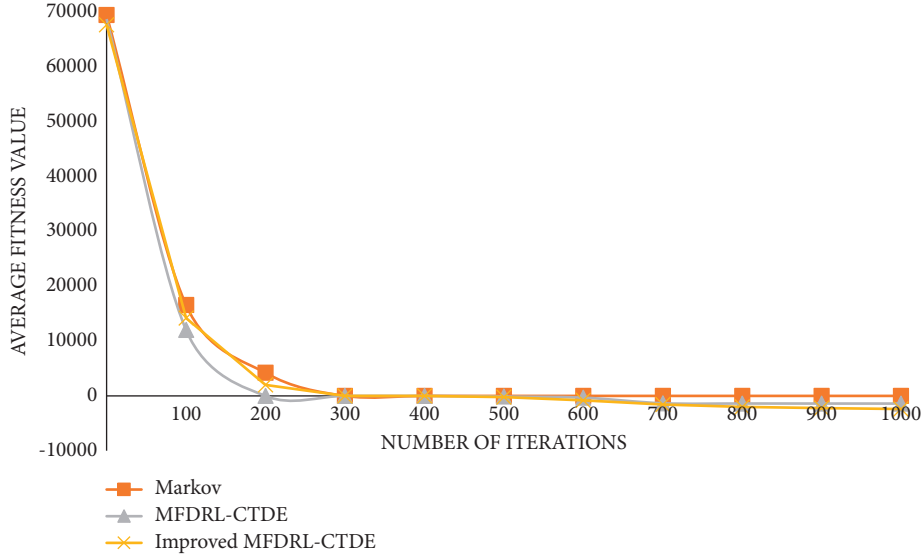


FIGURE 5: Graph of function 2.

priority experience playback, which can make the samples in the shared experience pool get priority rights, and the possibility of being sampled can be determined. At the same time, high-priority samples are selected many times, which leads to multiple playbacks, which will reduce the diversity of samples. At this time, important sampling weights are needed to solve this problem.

The formula for judging priority is as follows:

$$\delta_t^{ij} = r_t^{ij} + \gamma \max_{a'} Q^i(s_{t+1}^i, a'; \theta_{\text{tar}}, \theta_{\text{Vtar}}, \theta_{\text{Aatar}}) - Q^i(s_t^i, a_t^i; \theta_{\text{val}}^i, \theta_{\text{Vval}}^i, \theta_{\text{Aval}}^i). \quad (9)$$

Among them, when δ_t^{ij} is much greater than 0, it shows that the prediction accuracy has more room for improvement, which makes the convergence efficiency of the algorithm higher. δ_t^{ij} is time series difference error from the j -th sample of the i -th agent in the t -th iterative update, r_t^{ij} is an immediate reward from the j -th sample of the i -th agent in the t -th iterative update, $Q^i(s_{t+1}^i, a'; \theta_{\text{tar}}, \theta_{\text{Vtar}}, \theta_{\text{Aatar}})$ is the Q value obtained by the centralized target of the i -th agent, and $Q^i(s_t^i, a_t^i; \theta_{\text{val}}^i, \theta_{\text{Vval}}^i, \theta_{\text{Aval}}^i)$ is the Q value obtained from the estimation of the i -th agent.

The formula for calculating the probability G_{ij} that a sample is sampled is as follows:

$$G_{ij} = \frac{g_{ij}}{\sum_{k=1}^{N_g} g_{ik}}, \quad (10)$$

where N_g is the empirical pool capacity.

The formula for calculating sample priority g_{ij} is as follows:

$$g_{ij} = \left(|\delta_t^{ij}| + \sigma \right)^\alpha, \quad (11)$$

where σ is a smaller positive number, α is degree coefficient, which controls priority and has a value range of [0, 1], and when the value of degree coefficient is 1, it means that sampling is not based on priority, but on random sampling.

The formula of importance sampling weight w_{ij} is as follows:

$$w_{ij} = (N_g G_{ij})^{-\beta}, \quad (12)$$

where β is the correction degree parameter.

The formula of parameter training target value y_{ij} is as follows:

$$y_{ij} = \begin{cases} r_t^{ij}, & \text{End of current training round,} \\ r_t^{ij} + \gamma \max_{a'} Q^i(s_{t+1}^i, a'; \theta_{\text{tar}}, \theta_{\text{Vtar}}, \theta_{\text{Aatar}}), & \text{Others.} \end{cases} \quad (13)$$

The formula of the corrected loss function is as follows:

$$L(\theta_{\text{val}}^i, \theta_{\text{Vval}}^i, \theta_{\text{Aval}}^i) = \sum_j w_{ij} [y_{ij} - Q^i(s_t^i, a_t^i; \theta_{\text{val}}^i, \theta_{\text{Vval}}^i, \theta_{\text{Aval}}^i)]^2. \quad (14)$$

4. Simulation Experiment

4.1. Algorithm Testing. Six typical test functions are selected to calculate and compare the convergence rates of the three algorithms mentioned above. The first three are single-mode

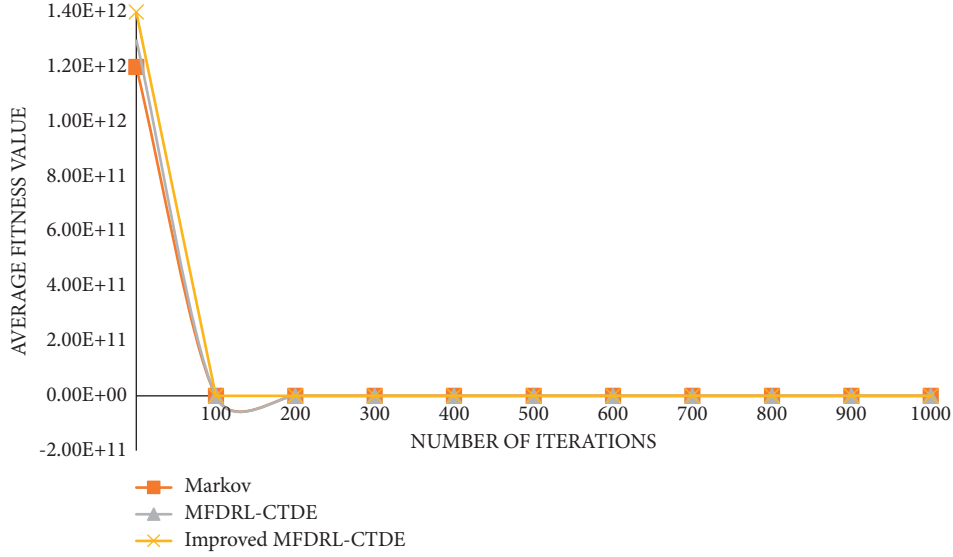


FIGURE 6: Graph of function 3.

benchmark test functions, and the latter three are multimode benchmark test functions. The formula for the test function is as follows:

$$f_1(x) = \sum_{i=1}^D x_i^2 - 100 \leq x_i \leq 100,$$

$$f_2(x) = \sum_{i=1}^D (|x_i + 0.5|)^2 - 100 \leq x_i \leq 100,$$

$$f_3(x) = \sum_{i=1}^D |x_i| + \prod_{i=1}^D |x_i| - 10 \leq x_i \leq 10,$$

$$f_4(x) = \sum_{i=1}^D (x_i^2 - 10 \cos(2\pi x_i) + 10) - 5.12 \leq x_i \leq 5.12,$$

$$f_5(x) = \frac{1}{400} \sum_{i=1}^D x_i^2 - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 - 300 \leq x_i \leq 300,$$

$$f_6(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i}\right) - \left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i)\right)$$

$$+ 20 + e - 32 \leq x_i \leq 32.$$

(15)

The average and standard deviation of the three algorithms are calculated after 1000 iterations according to the target test function. The significance of the calculation is that it can reflect the convergence speed of the algorithms and show whether the algorithms are stable or not. The calculation results are shown in Table 1.

The graph of function 1 is shown in Figure 4.

The graph of function 2 is shown in Figure 5.

The graph of function 3 is shown in Figure 6.

The graph of function 4 is shown in Figure 7.

The graph of function 5 is shown in Figure 8.

The graph of function 6 is shown in Figure 9.

It can be seen from the above comparison diagram that the stability and convergence speed of the three algorithms are very close in a single modal function. But, in the multimodal function, it is obvious that Markov does not show good stability, and the convergence rate is very slow. Compared with the improved MFDRL-CTDE algorithm and the common MFDRL-CTDE algorithm, the convergence speed of the former is slightly better than that of the latter, but, for the best performance position, the latter is obviously better than the former, and the latter is more stable. Therefore, in general, the improved MFDRL-CTDE algorithm is superior to the original algorithm.

4.2. Simulation Experiment

4.2.1. Objective Function. Set its standard aerobics action as action sequence P , set each person's aerobics action as a series of action sequence Q , compare them, and calculate the similarity value.

The formula of similarity evaluation is as follows:

$$\text{DIST}(P, Q) = \frac{\sum_{r=1}^R \text{dist}(K_{p_r}, K_{q_r})}{R}, \quad (16)$$

$$\text{Similarity}(P, Q) = \frac{1}{\text{DIST}(P, Q) + 1},$$

where R is the total number of key actions, K_{p_r} is the key action set in standard action sequence P , and K_{q_r} is the key action set in standard action sequence Q .

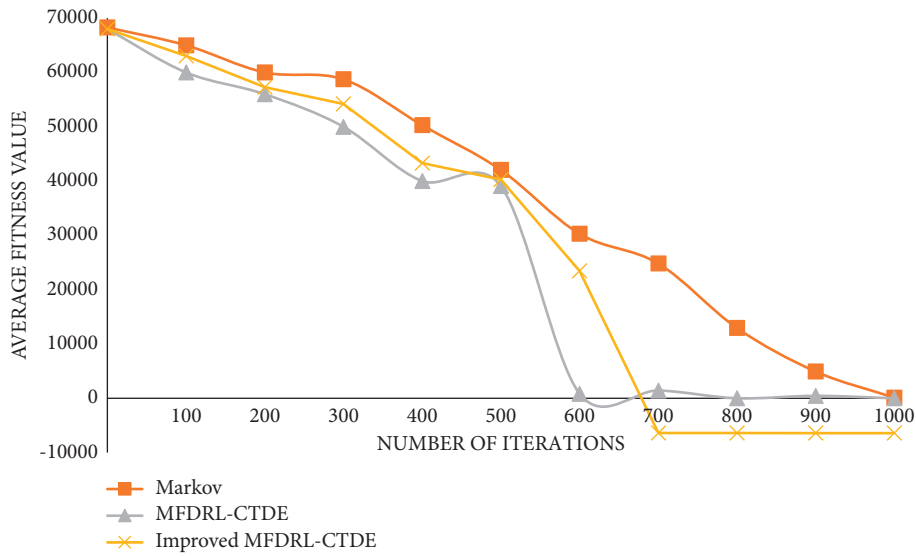


FIGURE 7: Graph of function 4.

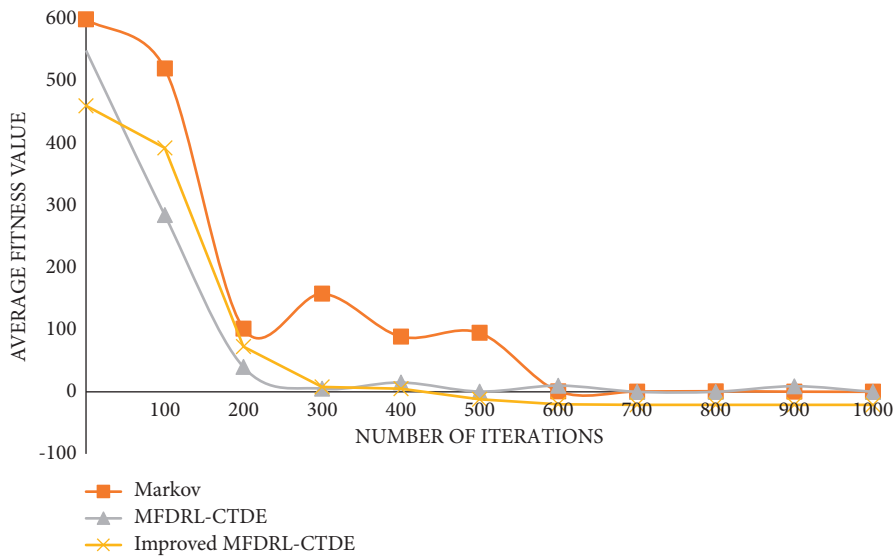


FIGURE 8: Graph of function 5.

4.2.2. *Simulation Experiment and Result Analysis.* In the aerobics competition, the movement of each member of the team is set into an action sequence, and the corresponding standard action sequence is compared with it. Suppose there are two aerobics teams participating in the five-person competition and the ten-person competition, respectively. The standard action sequence of aerobics displayed by each team is known. Under the same competition venue, each team completes the display action within 90 seconds at the same time, taking every 10 seconds as a judgment decision point, recording, and setting it into an action sequence to compare its similarity judgment decision.

By taking aerobics as an example, important actions are intercepted as one of the action sequences, and the picture of the intercepted actions is shown in Figure 10.

Put four intercepting actions into the action sequence, which are $a_t^1, a_t^2, a_t^3, a_t^4 \dots$, and compare the action sequences to calculate the similarity.

Set the volume of the shared experience pool to 2000, the number of playback samples to 10, $\bar{T} = 90, \gamma = 0, \alpha = 0.6, \beta = 0.4, \lambda = 40$.

The similarity comparison data of the five-person group is shown in Table 2.

The similarity comparison data of the ten-person group is shown in Table 3.

The five-person action sequence decision prediction pair is shown in Figure 11.

Therefore, for aerobics competition, the results of the three algorithms for the evaluation scheme decision of the competition action sequence are obvious. With the increase

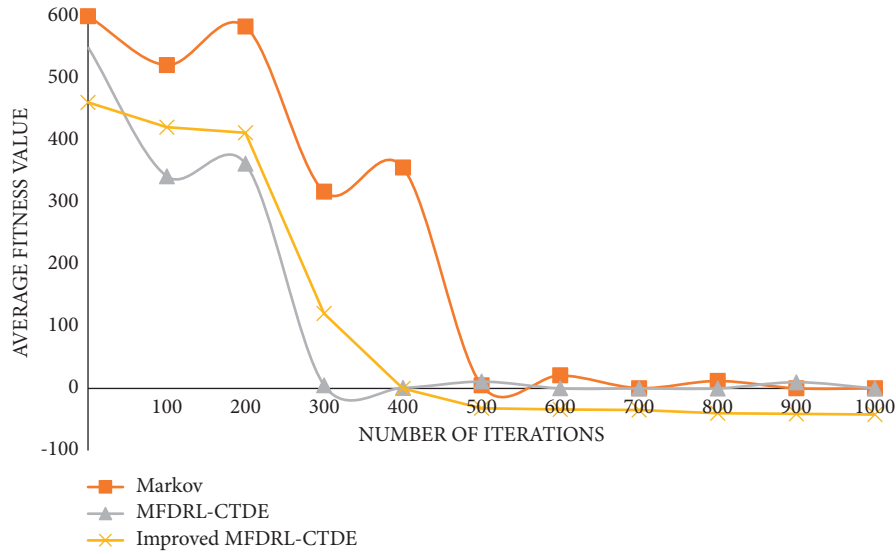


FIGURE 9: Graph of function 6.



FIGURE 10: Intercept action.

TABLE 2: Comparison data table of similarity of five-person group.

Time (S)	Actual value (%)	Markov's prediction value (%)	MFDRL-CTDE predictive value (%)	Improved MFDRL-CTDE prediction value (%)
10	80	70	78	80
20	77	68	86	80
30	81	71	75	79
40	85	73	80	83
50	84	69	85	86
60	80	72	75	81
70	79	65	75	79
80	82	70	80	83
90	86	70	81	84

of decision objects, the accuracy of Markov's algorithm is obviously reduced, which coincides with Markov's suitability for a single intelligent decision system. By comparing the MFDRL-CTDE algorithm and the improved MFDRL-CTDE algorithm, it can be clearly seen from the above figure that the improved MFDRL-CTDE algorithm is more stable for the accuracy of predicted values. Although there are very few cases where MFDRL-CTDE predicted values are more

accurate than the improved algorithm, most predicted values are still not accurate and stable enough. Therefore, it can be concluded that the improved MFDRL-CTDE algorithm is more stable and efficient for multi-intelligent decision-making systems. In the multi-intelligent decision-making, the system can make good use of all kinds of dance competition action contrast similarity, so that sports competition has a deeper step of development.

TABLE 3: Data table of similarity comparison of ten groups.

Time (S)	Actual value (%)	Markov's prediction value (%)	MFDR-CTDE predictive value (%)	Improved MFDR-CTDE prediction value (%)
10	75	60	73	71
20	73	61	70	71
30	80	59	85	77
40	78	65	80	80
50	81	68	86	80
60	76	62	70	75
70	80	56	75	81
80	83	60	80	83
90	86	64	80	85

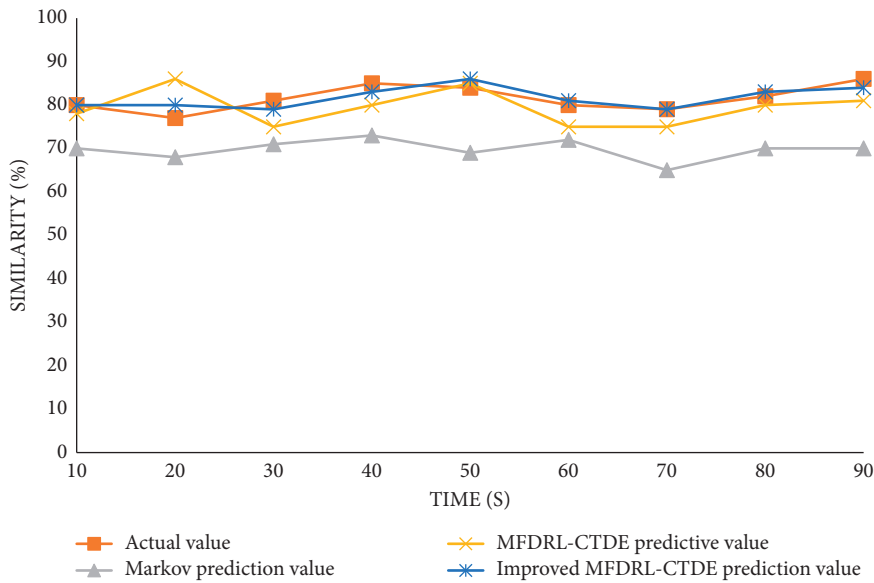


FIGURE 11: Comparison chart of decision prediction of five-person action sequence. The ten-person action sequence decision prediction pair is shown in Figure 12.

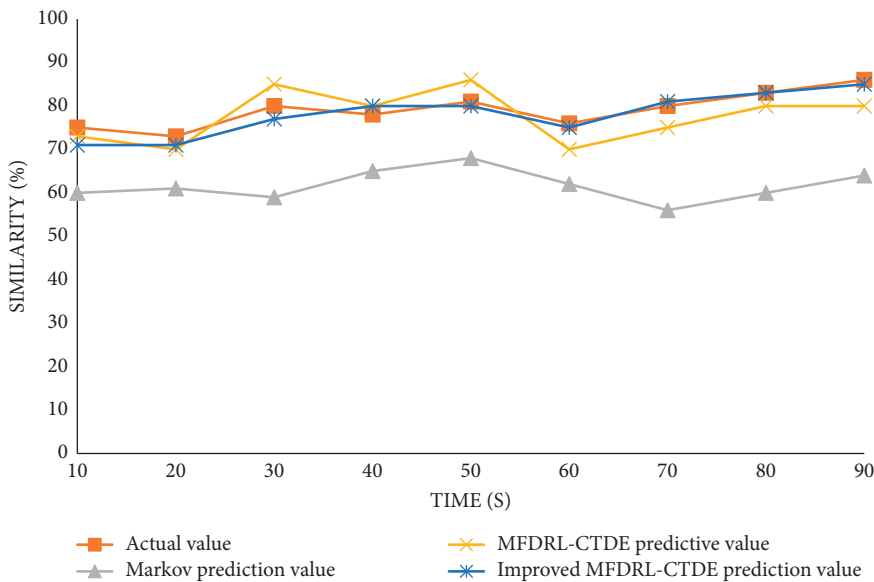


FIGURE 12: Comparison chart of decision prediction of ten-person action sequence.

5. Conclusion

Compared with MFDRL-CTDE, the improved MFDRL-CTDE algorithm is more suitable for multi-intelligent decision-making. Even after many iterations, the decision-making of the algorithm will not fall into an agent and then execute actions. Obviously, the actions are still executed according to priority, which ensures the efficiency of the algorithm. Markov is more suitable for single intelligent decision-making and cannot be effectively implemented for multi-intelligent decision-making. Therefore, for the scheme scheduling of aerobics competition, the improved MFDRL-CTDE has obvious advantages.

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 regarding this work.

References

- [1] L. Liu, Q. Li, and H. Song, "Intelligent Group Decision Support System for Cooperative Works Based on Multi-Agent System," in *Proceedings of the 10th International Conference on CSCW in Design, CSCWD 2006, May 3-5, 2006, Southeast University, Nanjing, China, IEEE, Nanjing, China, May 3-5 2006*.
- [2] X. Z. Hui and C. Wuhan, "Study on game theory in decision interaction for multi intelligent agents based on information fusion," *Computer Science*, vol. 40, no. 7, pp. 196–200, 2013.
- [3] J. Lu and Z. D. Ruan, "Intelligent multi criteria fuzzy group decision making for situation assessments," *Soft Computing-A Fusion of Foundations, Methodologies&Applications*, vol. 12, no. 3, 2008.
- [4] M. F Santos, F. Portela, and B. M Vilas, "INTCARE: multi agent approach for real time intelligent decision support in intensive medicine ICAART 2011," in *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence*, vol. 1, Artificial Intelligence, Rome, Italy, January 28-30, 2011.
- [5] K. G. Jolly, K. P. Ravindran, R. Vijayakumar, and R. Sreerama Kumar, "Intelligent decision making in multi," *Robotics and Autonomous Systems*, vol. 55, no. 7, pp. 589–596, 2007.
- [6] X. Sun, V. Gollnick, and Y. Li, "An intelligent multi criteria decision support system for systems design 10th AIAA aviation technology, integration, and operations (ATIO) conference (AIAA)," *Journal of Aircraft*, vol. 51, no. 1, pp. 216–225, 2010.
- [7] I. Ashikhmin and E. Furems, "UniComBOS-intelligent decision support system for multi criteria comparison and choice," *Journal of Multi-Criteria Decision Analysis*, vol. 13, no. 2-3, pp. 147–157, 2005.
- [8] D. L. Xu, G. Mccarthy, and J. B. Yang, "Intelligent decision system and its application in business innovation self assessment," *Decision Support Systems*, vol. 42, no. 2, pp. 664–673, 2006.
- [9] S. So and A. Bae, "introducing and development process in korea," *Journal of Korean Association of Physical Education Sport for Girls and Women*, vol. 15, no. 2, pp. 159–168, 2001.
- [10] A. Zhang, "Empirical Research on Reasonability of Competition Ratings of Aerobics Referees," in *Proceedings of the 3 International Symposium on Information Technology and Management Science*, American Research Press, Tianjin, China, September. 25 2011.
- [11] X. Li, "Study on the difficult movement of the sports aerobics world championships," *Contemporary Sports Technology*, vol. 3, no. 32, pp. 175–178, 2013.
- [12] H. Lu, "Check procedure for computer aided aerobics competition scoring," in *Proceedings of the Fourth International Symposium on Knowledge Acquisition and Modeling*, pp. 344–346, Sanya, China, October. 8-9 2011.
- [13] L. Guo, "The feature analysis of the step aerobics competition routine arrangement of the 26th FISU," *Advances in Information Sciences & Service Sciences*, vol. 5, Linfen City, China, 2013.
- [14] S. E. Wang, "An analysis on the characteristics of middle school students' aerobics competition in the 10th Guangdong Middle School Students' Games [J]," *Liaoning Sport Science and Technology*, vol. 35, no. 01, 2013.
- [15] A. H. Zhang, "From aerobics training and competitions to see the status of development of aerobics in China and countermeasures," *Advanced Materials Research*, vol. 926-930, pp. 4158–4161, 2014.

Research Article

Voice Recognition and Evaluation of Vocal Music Based on Neural Network

Xiaochen Wang¹ and Tao Wang² 

¹GuiZhou University of Finance and Economics, Guiyang, Guizhou 550000, China

²Beijing Technology and Business University, Beijing 100048, China

Correspondence should be addressed to Tao Wang; 2004030128@st.btbu.edu.cn

Received 3 March 2022; Revised 23 March 2022; Accepted 7 April 2022; Published 20 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiaochen Wang and Tao Wang. 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.

Artistic voice is the artistic life of professional voice users. In the process of selecting and cultivating artistic performing talents, the evaluation of voice even occupies a very important position. Therefore, an appropriate evaluation of the artistic voice is crucial. With the development of art education, how to scientifically evaluate artistic voice training methods and fairly select artistic voice talents is an urgent need for objective evaluation of artistic voice. The current evaluation methods for artistic voices are time-consuming, laborious, and highly subjective. In the objective evaluation of artistic voice, the selection of evaluation acoustic parameters is very important. Attempt to extract the average energy, average frequency error, and average range error of singing voice by using speech analysis technology as the objective evaluation acoustic parameters, use neural network method to objectively evaluate the singing quality of artistic voice, and compare with the subjective evaluation of senior professional teachers. In this paper, voice analysis technology is used to extract the first formant, third formant, fundamental frequency, sound range, fundamental frequency perturbation, first formant perturbation, third formant perturbation, and average energy of singing acoustic parameters. By using BP neural network methods, the quality of singing was evaluated objectively and compared with the subjective evaluation of senior vocal professional teachers. The results show that the BP neural network method can accurately and objectively evaluate the quality of singing voice by using the evaluation parameters, which is helpful in scientifically guiding the selection and training of artistic voice talents.

1. Introduction

Artistic voice mainly refers to the singing voices appearing in artistic performances on the stage, movies, television, and radio, the singing voices, lines, and other animal voice shapes and voice effects in operas. It uses the color and style of the sound to create different images to express the content of the work and reflect the colorful life; it is efficient and subtle in sound modeling, versatile, and profound; it is the artistic life of professional voice users, and in the selection of in the process of cultivating artistic performance talents, the evaluation of voice even occupies a very important position. Therefore, an appropriate evaluation of the artistic voice is crucial [1–5].

Objective evaluation of artistic voice is an important part of medical research on artistic voice. Using microcomputer and acoustic knowledge to analyze artistic voice signal is an

effective method to evaluate pronunciation. It has the advantages of noninvasiveness and objective and comparable data [6, 7]. Quantitative analysis and objective evaluation of various factors and laws in the pronunciation of singing artistic voices can help solve problems such as the selection of reasonable artistic voices, the evaluation and improvement of voice training methods, and the diagnosis and treatment of voice diseases.

At present, the most commonly used evaluation method in artistic voice evaluation is subjective listening perception evaluation, but the problems it faces cannot be ignored. For example, when selecting and cultivating artistic performance talents, the evaluation of vocal talents is basically based on the voice experts using their own experience to listen to and feel the sound quality of the test subjects as a whole and use simple scores or comments to represent the evaluation

results. Due to the differences in the evaluation standards among the evaluators and the influence of various factors, the evaluation results are lacking in objectivity, accuracy, and fairness to a certain extent. Moreover, the results of subjective listening perception evaluation often cannot fully reflect the vocal condition of vocal talents, and it is difficult to provide an effective reference for various problems in vocal music teaching. Although some schools use voice medical examination as one of the reference basis for selecting artistic performing talents, the medical examination process will undoubtedly bring psychological and physical pain to candidates. In addition, since the evaluation of artistic voice runs through the whole process of talent selection, training, diagnosis, and treatment of voice problems, with the continuous increase of art students, the workload of voice evaluation has increased exponentially, thus affecting the human and material resources. Demand has also increased substantially, putting the evaluation of artistic voices in trouble. In order to solve the problems existing in the evaluation of subjective auditory perception, and to make the evaluation results more accurately reflect the functional state of people's voice during vocalization, it is necessary to find a new evaluation method, which takes the physical characteristics of the auditory as the standard and is not subject to subjective factors. With the rapid development of electronic technology and computer technology, the interaction between computers and other electronic products and human beings has become more and more extensive and in-depth [8–14]. They penetrate into almost every aspect of human life, give human beings great help, and also work for the voice. The speaker brought the gospel and made the objective evaluation of the artistic voice possible.

The objective evaluation of artistic voice not only plays a good guiding role in the scientific selection and training of singing and performing talents, but also plays a role in studying the vocal pathology of speakers and preventing and treating voice diseases. If you want the vocal music industry to become more and more prosperous, you must pay attention to the training and delivery of vocal talents. Singers, voice actors, performers, and vocal teachers need to protect their voices from further damage. Therefore, the correct objective evaluation of artistic voice is crucial. Researchers using computer and acoustics knowledge to explore objective and scientific evaluation methods for vocal quality in singing art have achieved initial results. Because of the noninvasive and excellent computing power of computers, its evaluation results are objective, reliable, accurate, and efficient. The purpose of objective evaluation of the voice includes the following aspects: (1) to increase the interest in singing; (2) to protect the tender voice; (3) to improve artistic accomplishment; (4) to evaluate the scope and extent of the disease; (5) to try to diagnose the disease; (6) after predicting and correcting the vocalization problem [15–20].

In the research of establishing the objective evaluation method of singing artistic voice, it is mainly divided into the evaluation method based on neural network and the evaluation method based on feature matching. The evaluation method based on neural network refers to the use of RBF network, BP neural network, CNN network, etc., to

automatically extract discriminative features from the sound signal, and the use of softmax, support vector machine, and other classifiers to classify and discriminate the sound quality. It is a rating or score for an objective assessment of the quality of a singing voice. Neural networks have developed rapidly in recent years due to their excellent feature classification capabilities and the ability to stack multiple layers. Currently, the most advanced deep learning methods in the world are based on neural networks. Some people use the neural network method to build a model network for objective evaluation of artistic voices. The reasons for this are as follows: (1) from a probabilistic point of view, RBF does not have the same good probability characteristics as Softmax. (2) With the increase of the number of network layers, both the network structure and the complexity of the weight parameters that need to be trained increase accordingly, and the neural network has three prominent features such as local connection, weight sharing, and pooling operation, so the network structure is simple, the weight parameters are few, the training effect is good, and the classification accuracy rate is high. (3) The voice audio sample is an evaluation model of a one-dimensional signal neural network, which can reflect its characteristics more objectively. With the development of science and technology, the method of evaluating the quality of artistic voice based on artificial intelligence technology has been widely recognized by people in the vocal music industry. Therefore, it is theoretically and technically feasible to apply the neural network method to the objective evaluation of artistic voice in this paper [21–24].

To sum up, the objective evaluation of singing artistic voice is very meaningful for the selection and cultivation of singing artistic talents. It can not only select high-quality voices but also help improve the training methods of voices, so that singers can speak scientifically. More importantly, it allows singers to prevent related voice diseases, keep their voices young, and prolong their artistic life.

Therefore, using a computer to analyze the voice signal can effectively evaluate the quality of the sound quality, inability to carry out objective and fair evaluation of each voice sample and other defects. This paper analyzes the singing voice signal, extracts the characteristic parameters that characterize the singing voice, and uses the BP neural network to evaluate it objectively through machine learning and compare the evaluation results with the subjective evaluation [25, 26].

2. Acoustic Parameters

Singing is the language of art, the supplement and development of language. Artistic voice requires not only clear sentences, but also better sound quality. Because the core of this indicator is the voice, it is stipulated that the voice must be professional and artistic. A good artistic voice is characterized by penetrating singing, which can freely control the strength of breath and the level of energy. In addition, it should also have certain tension and endurance. The difference between an artistic voice and an ordinary voice is that the vocal cords of the former can freely control

continuous vibration within a specified period of time, such as continuous vibration when singing. All in all, artistic voices exhibit many unique characteristics acoustically. Because acoustics is a discipline that studies the entire process of sound from generation, propagation to feedback, therefore, this is also the basis for studying scientific voices and screening qualified artistic voices.

The objective evaluation of singing is realized by computer programming in Matlab, and the first formant, third formant, fundamental frequency, sound range, fundamental frequency perturbation, first formant perturbation, third formant perturbation, and average energy of singing voice are extracted as evaluation parameters.

2.1. Formant Extraction. Formant is a theory that studies the resonance and quality of voice. In general, the lower two of these peaks, the first and second formants, basically define the vowel timbre of the sound, while the higher third, fourth, and fifth formants affect the personal character of the sound and the musical timbre. The long-term research results of domestic and foreign scholars on vocal formants in singing art show that the first and third formants are important objective reference data to measure the vocal technical level of singers. Therefore, the experiment selects the first and third formants as objective evaluation parameters and adopts the AR model peak detection method to extract the first and third formants. The formant time is plotted in Figure 1.

2.2. Fundamental Frequency Extraction. The fundamental frequency refers to the fundamental frequency at which the vocal cords vibrate, also known as the fundamental frequency. In addition to good timbre, an excellent artistic voice should also have a certain tension, that is, whether a certain sound can reach a certain height. Fundamental frequency is one of the important parameters for objective evaluation of artistic voice. There are many methods to extract fundamental frequency, such as autocorrelation function, average amplitude difference function, autocorrelation, and wavelet transform. After calculation and comparison, this experiment adopts the improved algorithm to extract the pitch period, so as to calculate the fundamental frequency. That is, the pitch period detection algorithm of artistic voice based on improved wavelet change mentioned later. The experimental steps are as follows:

- (1) Preprocess the original noisy signal, extract useful information segments from the processed signals, and combine these information segments into artistic voice segments.
- (2) Arrange the extracted artistic voice segments and filter out the DC components in them; then, set the signal-to-noise ratio to superimpose the noise; finally, use the DWT wavelet transform, and use the obtained low-frequency coefficients to reconstruct the target signal.
- (3) Perform pitch detection on the reconstructed artistic voice segment.

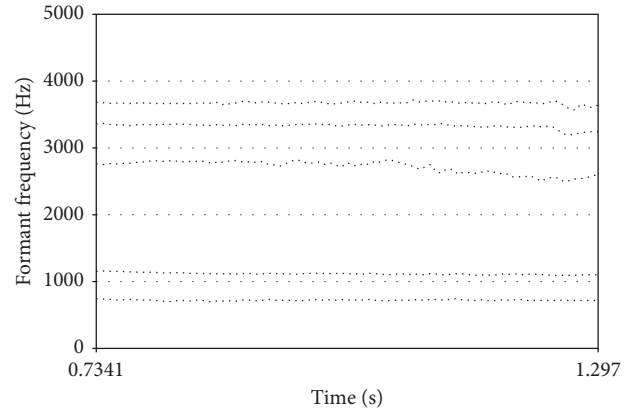


FIGURE 1: Formant time plot.

2.3. Sound Field Extraction Method. The simplest method for estimating the pitch range is to select the maximum and minimum pitch values that appear in the singing voice and musical score. However, there is much pitch data in the singing voice and musical score, and only two of them are selected, which is unavoidable by chance, and it is necessary to use statistical methods. The method of estimating the sound range in this experiment is to take the average and standard deviation of the D values of all pitches in the singing voice and score. The corrected autocorrelation waveforms before and after center clipping are compared in Figure 2.

Pitch is determined by the vibration frequency of the object (vocal cord). That is, if the number of vibrations in a certain unit time is large, the pitch will be high; if the number of vibrations is small, the sound will be low. Lin Tao and others pointed out in "Beijing Voice Experiment Record": "Tone and intonation are both expressions of pitch. After all, pitch is not the direct reality of the fundamental frequency, but an abstraction cut across the pitch curve based on Hertz's things". Therefore, the pitch is described by the D value in the experiment, which is the logarithmic scale of pitch and is defined as

$$D = 12 * \log_2 \left(\frac{F}{F_0} \right). \quad (1)$$

That is, it is the step difference of the pitch F in Hertz with respect to the reference frequency F_0 . The D value is a dimensionless number, and its unit can be taken as "degree," which is replaced by D for the convenience of writing.

Take the average and standard deviation of all pitch D values in the singing voice and score:

$$\bar{D} = \frac{1}{N} \sum_{j=1}^N D_j. \quad (2)$$

The standard deviation σ is

$$\sigma = \sqrt{E \left[(D_j - \bar{D})^2 \right]}, \quad (3)$$

where E represents the average, and N is the number of pitch data elements. The domain width S is

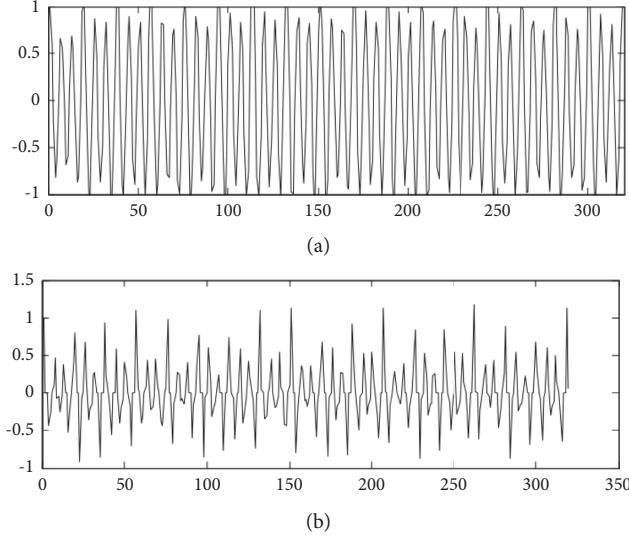


FIGURE 2: Corrected autocorrelation waveforms before and after center clipping.

$$S = 4\sigma. \quad (4)$$

The fundamental frequency perturbation refers to the rate of change of the fundamental frequency of the sound wave between adjacent cycles, which is used to measure the difference between a specified cycle and the adjacent previous or next cycles, reflecting the frequency difference between the vocal cord vibration cycles. Mathematical definition of fundamental frequency perturbation:

$$Jitter = \frac{1}{N-1} \sum_{i=1}^N \left| \frac{1}{F_{0i}} - \frac{1}{F_{0(i-1)}} \right|. \quad (5)$$

In the formula, F represents the fundamental frequency of the i th cycle.

The first and third formant perturbations measure the rate of change of the first and third formants between adjacent periods, respectively. The mathematical definitions of formant perturbation are as follows:

$$\begin{aligned} & \frac{1}{N-1} \sum_{i=1}^N \left| \frac{1}{F_{1i}} - \frac{1}{F_{1(i-1)}} \right|, \\ & \frac{1}{N-1} \sum_{i=1}^N \left| \frac{1}{F_{3i}} - \frac{1}{F_{3(i-1)}} \right|. \end{aligned} \quad (6)$$

The average energy represents the relative magnitude of the singing voice signal in the same environment. Mathematical definition of short-term energy of speech signal:

$$E_n = \sum_{k=-\infty}^{+\infty} x^2(k)w(n-k). \quad (7)$$

3. Neural Networks

Since the neural network is formed by imitating the biological neural network of the human brain, it is very close to the operating rules of the human brain when processing

information and has strong self-learning and adaptive capabilities. The network automatically adjusts its structure according to different input samples. Parameters to give the desired output: because the knowledge is stored in the connection weights, the neural network can realize various nonlinear mappings by adjusting the weights; due to the relative independence between neurons in each layer, when the neural network stores information, it is distributed, so its robustness and fault tolerance are good; in addition, the neural network is composed of a large number of simple processing units connected in parallel, making it extremely easy to implement in hardware and running extremely fast.

These characteristics of neural network make it have a very good performance in many fields, and it has played an important role in many industries such as aerospace, automotive, manufacturing, finance, and telecommunications. The characteristics of neural network make it especially suitable for speech signal processing. Since 1980, the application of neural network in speech signal processing has been very active, and the more prominent one is speech recognition. This paper attempts to use the neural network method to evaluate the level of artistic voice signals.

A neural network is an information processing system or mathematical model that imitates the function of the human brain, and there are many different network models. Since the objective evaluation of singing needs to determine the weights (weight coefficients) of each evaluation parameter, the backpropagation network model, often called the BP network model, is selected. Through the reverse adjustment of the weights in the network learning process, each weight can be obtained, a set of optimal solutions. BP neural network includes input layer, hidden layer, and output layer. Using BP neural network to objectively evaluate singing includes two processes, training process and evaluation process, the training process is based on the evaluation parameters of the training samples and the corresponding subjective evaluation scores through neural network learning and training to establish an evaluation model; the evaluation process uses the

evaluation parameters of the samples to be evaluated. Input the established evaluation model, get the objective evaluation score, and compare it with the corresponding subjective evaluation score to test the usability of the established evaluation model (see Figure 3).

The BP neural network adopts the error back propagation learning algorithm, which is based on the Delta learning rule and uses the gradient search technology to minimize the mean square error between the actual output and the expected output of the network. The process of network learning is a process of correcting weights while propagating backwards. In such a network, the learning process consists of forward propagation and back-propagation. In the forward process, the input signal is processed layer by layer from the input layer through the hidden layer unit and transmitted to the output layer. The state of each layer of neurons only affects the state of the next layer of neurons. If the expected output cannot be obtained at the output layer, turn to backpropagation and return the output error according to the original connection path. By modifying the weights of neurons in each layer, the error signal is minimized. Once the appropriate network connection values are obtained, the new samples can be non-linearly mapped.

Use S_j to calculate the output b_j of each intermediate unit through the transfer function:

$$S_j = \sum_{i=1}^n w_{ij} a_i - \theta_j, \quad (8)$$

$$b_j = f(s_j).$$

Use the output b_j of the intermediate layer, the connection weight v_{jt} and the threshold γ to calculate the output L_t of each unit of the output layer, and then use the transfer function to calculate the response C_t of each unit of the output layer.

$$L_t = \sum_{j=1}^p v_{jt} b_j - \gamma_t, \quad (9)$$

$$C_t = f(L_t).$$

Theoretically, it has been shown that a network with bias and at least one sigmoid hidden layer plus a linear output layer can approximate any rational number. Increasing the number of layers can further reduce the error and improve the accuracy, but it also complicates the network, thereby increasing the training time of the network weights. The improvement of error accuracy can actually be obtained by increasing the number of neurons, and its training effect is easier to observe and adjust than increasing the number of layers. In general, priority should be given to increasing the number of neurons in the hidden layer.

The improvement of network training accuracy can be obtained by using a hidden layer and increasing the number of neurons. In terms of structure implementation, it is much simpler than increasing the number of hidden layers. How many hidden layer nodes to select is appropriate, and there is no clear regulation in theory. In the specific design, the more practical way is to compare the training of different numbers

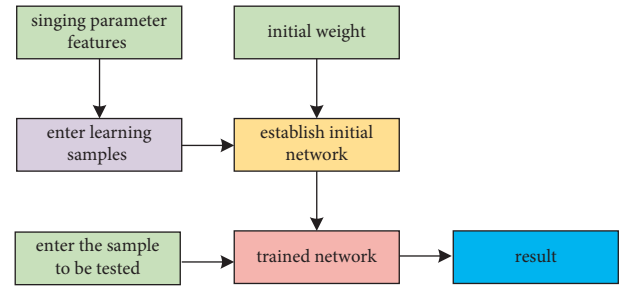


FIGURE 3: The established evaluation model.

of neurons and then add a little margin appropriately. More broadly, the number of neurons in the hidden layer = 2 * the number of input layers: 4 * the number of input layers. Since the system is nonlinear, the initial value is very important to whether the learning reaches a local minimum, whether it can converge, and the length of the training time. If the initial value is too large, the weighted input sum n falls into the saturation region of the s-shaped activation function, causing its derivative to be very small, which makes the revised weights close to zero, and the adjustment process almost stops. Therefore, it is generally hoped that the output value of each neuron after initial weighting is close to zero, which ensures that the weights of each neuron can be adjusted where their sigmoid activation function changes the most. Therefore, a random number with an initial weight between $(-1, 1)$ is generally taken.

The learning rate determines the amount of weight change produced in each loop training. A large learning rate may lead to instability of the system; but a small learning rate leads to a longer training time and may converge very slowly, but it can ensure that the error value of the network will not jump out of the trough of the error surface and eventually tend to the minimum error value. Therefore, in general, we tend to choose a smaller learning rate to ensure the stability of the system. The learning rate is selected in the range of 0.01–0.8. The prediction comparison is shown in Figure 4.

4. Simulation Experiments

Although subjective auditory perception evaluation is affected by subjective factors and has problems such as low accuracy and stability, it is undeniable that subjective auditory perception evaluation is the only reference standard for testing and evaluating the effectiveness of objective test parameters of voice and vocal function inspection. In order to provide a supervisory signal to the input training samples of the neural network and verify the effectiveness of the objective evaluation method in this paper, this paper conducts subjective listening perception evaluation of artistic voice signal. In order to verify the practicability of the objective evaluation method in this paper, two vocal music teachers and two vocal music graduate students were invited as judges to listen to 22 singing recording materials separately and scored on a 10-point scale, and the average score was taken, and the score was greater than 6 points. A score of less than 6 is considered poor. Put the subjective evaluation results together with the fundamental frequency, F1, F3,

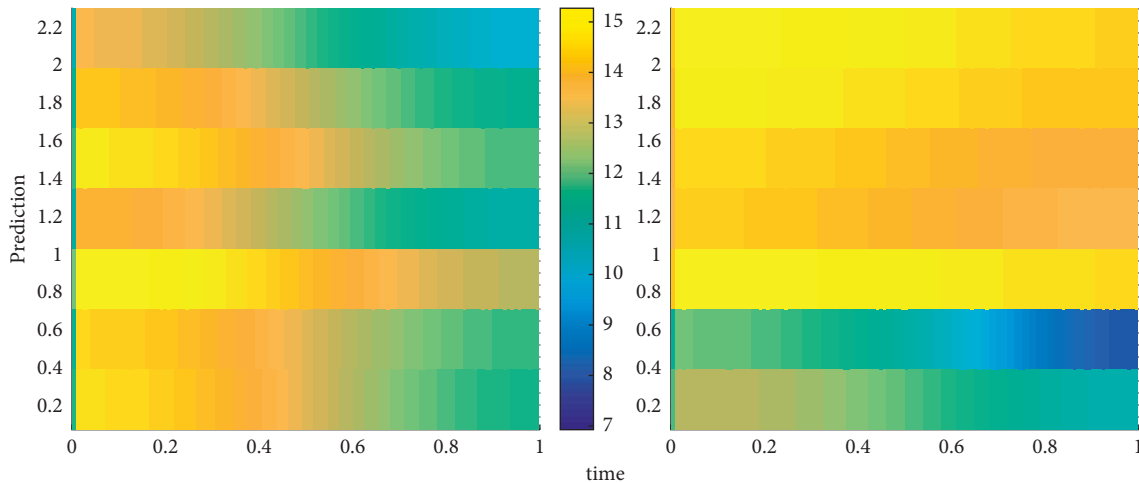


FIGURE 4: Prediction comparison.

extracted from the corresponding samples to prepare for the experiment.

The objective evaluation process includes training process and evaluation process. In the training process, the evaluation parameters of the training samples and their subjective evaluation results are learned through a neural network to obtain a trained network; in the evaluation process, the evaluation parameters of the remaining samples to be tested are input into the trained network to obtain objective evaluation results. The objective evaluation results were compared with the corresponding subjective evaluation results to test the accuracy of the neural network artistic voice objective evaluation model after training on the voice evaluation results. Neural network is a nonlinear system, which can fully approximate nonlinear systems of arbitrary complexity and has the ability of self-organization, self-adaptation, and self-learning for information processing. Error Backpropagation (BP) neural network is a multilayer feed-forward neural network with continuous transfer function. The training method is the error backpropagation algorithm, and the weights and thresholds of the network are constantly modified with the goal of minimizing the mean square error. Finally, the high-precision fitting of the data is carried out. In this paper, the BP neural network is used to objectively evaluate the artistic voice. The objective evaluation process of artistic voice includes training process and evaluation process. The training process will randomly select the evaluation parameters of the training samples and their subjective evaluation results through the BP neural network to learn to obtain the trained network; the evaluation process will input the evaluation parameters of the remaining samples to be tested into the trained network to obtain the objective evaluation results, which are combined with the results. The corresponding subjective evaluation results are compared to test the correctness of the voice evaluation results of the trained BP neural network. The error is plotted in Figure 5.

The specific process selects the first formant, the third formant, the fundamental frequency, the pitch range, the fundamental frequency perturbation, the first formant perturbation, the third formant perturbation, and the

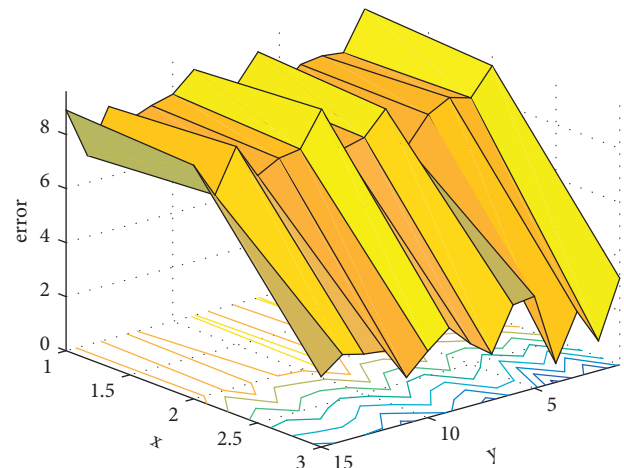


FIGURE 5: Error.

average energy of the training sample singing voice as evaluation parameters and normalizes and eliminates the effects of different orders of magnitude that are used as input parameters, and the scores scored by senior professional teachers in subjective evaluation of singing quality are used as output parameters to establish a nonlinear relationship model between training sample singing quality scores and evaluation parameters. After repeating the training, it is found that the neural network with the number of neurons in the hidden layer of 25 has the smallest approximation error to the singing training samples, so the number of neurons in the hidden layer is selected as 25, and the transfer functions of the hidden layer and the output layer are selected as tansig and logsig, respectively. The initial values of the weights and thresholds are random numbers between $[0, 1]$, and the maximum training error is set to 0.001. The optimal connection weights and thresholds are obtained through the training and learning of the neural network, and the neural network evaluation model is obtained. Finally, the parameters of the singing voice samples to be evaluated are used as input parameters, and the score of the singing voice is predicted by the trained neural network model.

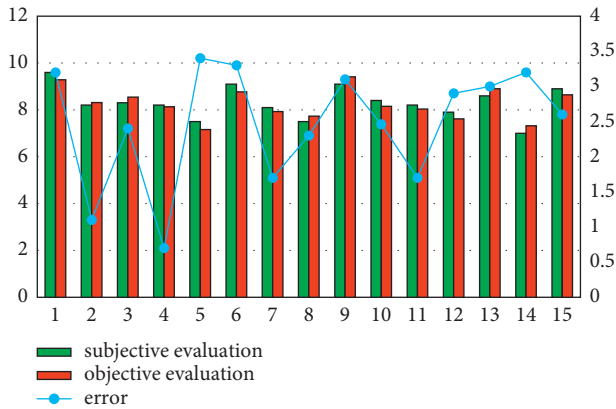


FIGURE 6: Differences between the objective evaluation and the subjective evaluation.

Figure 6 shows the differences between the objective evaluation scores and the subjective evaluation scores. Comparing the error between the objective evaluation scores and the subjective evaluation scores of senior professional teachers, the neural network method is within 3.4%. Since the number of high scores and low scores in the singing samples is small, and this is also the case for the randomly selected training set, the error between high scores and low scores is relatively large in objective evaluation.

Randomly select 6 good and bad singing signals from the subjective evaluation results, as the input of training samples, and get the trained network after training. Input the remaining 10 singing samples to be tested into the trained network to obtain objective evaluation results. When the final iteration number is 2000, the mean square error between the expected output and the actual output is calculated to be 0.1891, as shown in Figure 7. At this time, compared with subjective evaluation, the correct rate of objective evaluation reaches 90%. The distribution of the 12 samples involved in training is shown in Figure 8.

From the perspective of data normalization, since the BP neural network uses the idea of gradient descent to adjust the parameters of the network, the range of the input data will affect the update of the gradient value, influencing and improving the accuracy of the network evaluation results, and the input acoustic parameters need to be normalized, so the training time of the BP neural network is relatively long; the improved neural network is developed from the associative learning theory and competitive algorithms. As long as the distance between the input vector and the competition layer is directly calculated, there is no need to normalize and orthogonalize the data, so it is simpler and easier to realize the objective evaluation of artistic voice.

From the perspective of network training, the choice of the number of neurons in the hidden layer/competitive layer has a great impact on the performance of the network. Generally, a larger number of neurons in the hidden layer/competitive layer can bring better performance. However, the network structure will be more complex, resulting in too long training time. In the training of the BP neural network, the number of neurons in the hidden layer is selected to be 14 to obtain the smallest training error, while the number of

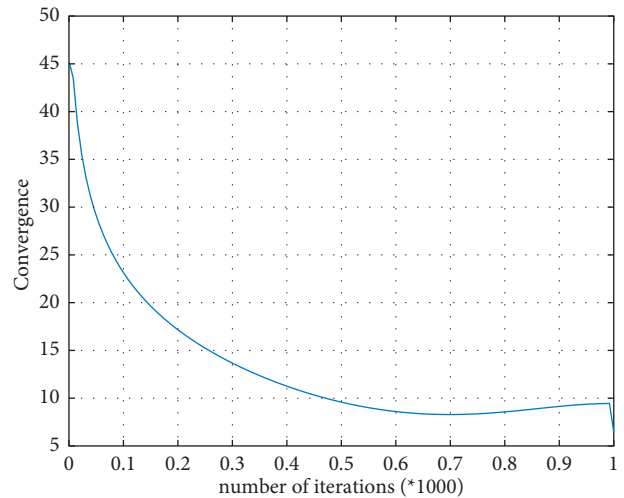


FIGURE 7: Convergence.

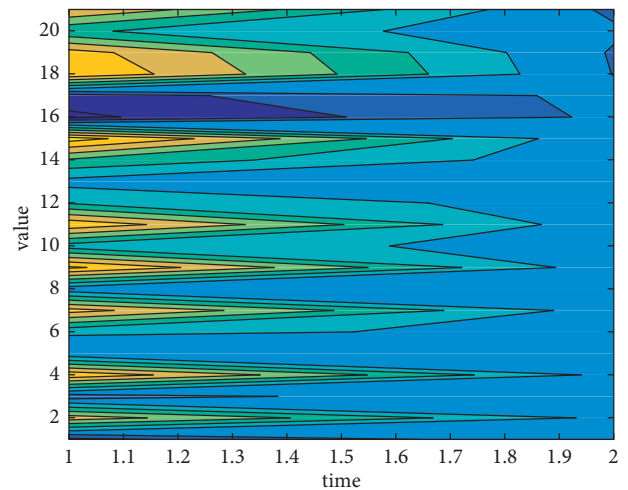


FIGURE 8: Evaluated value.

neurons in the competition layer of the improved neural network is set to 7. The minimum training error can be achieved. The number of training steps also affects the training time of the network. The more the training steps, the longer the training time. Experiments show that when the training errors of the BP neural network and the improved neural network reach the minimum, the training time is 10 seconds and 8 seconds, respectively. From the evaluation results, the use of BP neural network and improved neural network to objectively evaluate artistic voice signals can achieve satisfactory results. The accuracy of the BP method is 85.2%, and the accuracy of the improved method is 88.9%.

5. Conclusion

In this paper, the first formant, the third formant, the fundamental frequency, the pitch range, the fundamental frequency perturbation, the first formant perturbation, the third formant perturbation, and the average energy of the singing sample are extracted by speech analysis technology as evaluation parameters. The network method objectively

evaluates the singing quality, and the results show that the objective evaluation is consistent with the subjective evaluation of senior professional teachers. Because artistic voice singing is affected by multiple factors, and there is a complex mutual influence relationship between each factor, it has a highly uncertain nonlinear relationship. The neural network method considers the nonlinear mapping relationship between them and evaluates the singing voice. The quality is accurate. The experimental results show that the multiparameter artistic voice objective evaluation model established by the training samples has good consistency with the subjective auditory evaluation results and can comprehensively reflect the physiological function of the larynx.

Based on the above experimental results, the objective evaluation of singing artistic voice needs to be further improved and developed as follows: gender, vocal part increases male bass, middle, soprano and female bass, and middle and soprano singers. The evaluation model is established by gender, so as to more comprehensively explore the quantitative impact of various acoustic characteristic parameters on the sound quality, so as to allow singers to better grasp the vocalization laws and characteristics, and to more scientifically adjust the use of each vocalization organ.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] M. N. Anh, D. X. Bien, and L. Fortuna, "Voice recognition and inverse kinematics control for a redundant manipulator based on a multilayer artificial intelligence network," *Journal of Robotics*, vol. 2021, pp. 232–243, 2021.
- [2] S. LevAri, "EXPRESS: People with larger social networks show poorer voice recognition," *Quarterly Journal of Experimental Psychology* (2006), vol. 75, no. 3, Article ID 17470218211030798, 2021.
- [3] H. C. Hu, S. Y. Chang, C. H. Wang et al., "Deep learning application for vocal fold disease prediction through voice recognition: preliminary development study," *Journal of Medical Internet Research*, vol. 23, no. 6, Article ID e25247, 2021.
- [4] E. Holmes, G. To, and I. S. Johnsrude, "How long does it take for a voice to become familiar? Speech intelligibility and voice recognition are differentially sensitive to voice training," *Psychological Science*, vol. 32, no. 6, Article ID 956797621991137, 2021.
- [5] J. J. Ye, M. R. Tan, and C. H. Shum, "Using timestamp data to assess the impact of voice recognition on the efficiency of grossing biopsies," *Archives of Pathology & Laboratory Medicine*, vol. 145, no. 5, pp. 599–606, 2021.
- [6] C. Li, G. Jiang, and J.-M. Dewaele, "Understanding Chinese high school students' foreign language enjoyment: Validation of the Chinese version of the foreign language enjoyment scale," *System*, vol. 76, no. 76, pp. 183–196, 2018.
- [7] Y. Zhu and Y. Jiang, "Optimization of face recognition algorithm based on deep learning multi feature fusion driven by big data," *Image and Vision Computing*, vol. 104, no. 1, Article ID 104023, 2020.
- [8] O. Murton, R. Hillman, and D. Mehta, "Cepstral peak prominence values for clinical voice evaluation," *American Journal of Speech-Language Pathology*, vol. 29, no. 3, pp. 1596–1607, 2020.
- [9] I. Corrêa Nunes, E. Kohlbeck, F. Hãnsch Beuren, A. Borges Fagundes, and D. Pereira, "Life cycle analysis of electronic products for a product-service system," *Journal of Cleaner Production*, vol. 314, Article ID 127926, 2021.
- [10] M. Lazoryszczak, "Spectral features in the singing voice evaluation using simulated imperfections," *Procedia Computer Science*, vol. 176, pp. 1366–1373, 2020.
- [11] Z. Deng, S. Song, and H. Tan, "A new analytical framework for urban size and policies," *Applied Economics Letters*, vol. 28, no. 18, pp. 1610–1619, 2021.
- [12] C. C. Carvalho, D. M. Silva, A. D. de Carvalho Junior et al., "Pre-operative voice evaluation as a hypothetical predictor of difficult laryngoscopy," *Anaesthesia*, vol. 74, no. 9, pp. 1147–1152, 2019.
- [13] J. O. Kim, "Meta-analysis of semi-occluded vocal tract exercise studies on subjective voice evaluation," *Journal of speech-language & hearing disorders*, vol. 28, no. 2, pp. 1–11, 2019.
- [14] Z. Yuhui, C. Mengyao, C. Yuefen, L. Zhaoqian, and L. Yao, "An automatic recognition method of fruits and vegetables based on depthwise separable convolution neural network," *Journal of Physics: Conference Series*, vol. 1871, no. 1, pp. 1–9, 2021.
- [15] B. Jaber, R. Remman, and N. Matar, "Repetitive voice evaluation in dysphonic teachers: office versus home," *Journal of Voice*, vol. 34, no. 5, pp. 675–681, 2019.
- [16] B. You, H. Qi, L. Ding et al., "Fast neural network control of a pseudo-driven wheel on deformable terrain," *Mechanical Systems and Signal Processing*, vol. 152, Article ID 107478, 2021.
- [17] Y.-D. Zhang, S. C. Satapathy, D. S. Guttery, J. M. Górriz, and S.-H. Wang, "Improved breast cancer classification through combining graph convolutional network and convolutional neural network," *Information Processing & Management*, vol. 58, no. 2, Article ID 102439, 2021.
- [18] Z. Zhou, A. Gao, W. Wu et al., "Parameter estimation of the homodyned K distribution based on an artificial neural network for ultrasound tissue characterization," *Ultrasonics*, vol. 111, Article ID 106308, 2021.
- [19] P. Wang, K. Gu, J. Hou, and B. Dou, "An automatic recognition method for airflow field structures of convective systems based on single Doppler radar data," *Atmosphere*, vol. 11, no. 2, pp. 211–232, 2020.
- [20] J. Zhang, R. Jiang, B. Li, and N. Xu, "An automatic recognition method of microseismic signals based on EEMD-SVD and ELM," *Computers & Geosciences*, vol. 133, no. C, Article ID 104318, 2019.
- [21] R. Arora, B. Raman, K. Nayyar, and R. Awasthi, "Automated skin lesion segmentation using attention-based deep convolutional neural network," *Biomedical Signal Processing and Control*, vol. 65, Article ID 102358, 2021.
- [22] B. Ghorbani, A. Arulrajah, G. Narsilio, S. Horpibulsuk, and M. W. Bo, "Shakedown analysis of PET blends with demolition waste as pavement base/subbase materials using experimental and neural network methods," *Transportation Geotechnics*, vol. 27, Article ID 100481, 2021.

- [23] Y. Bai, J. Xie, C. Liu, Y. Tao, B. Zeng, and C. Li, "Regression modeling for enterprise electricity consumption: a comparison of recurrent neural network and its variants," *International Journal of Electrical Power & Energy Systems*, vol. 126, no. PA, Article ID 106612, 2021.
- [24] Z. Zhao, B. Li, X. Kang, J. Lu, and T. Liu, "Phase unwrapping method for point diffraction interferometer based on residual auto encoder neural network," *Optics and Lasers in Engineering*, vol. 138, Article ID 106405, 2021.
- [25] C. Mai, "A study on the current situation and countermeasures of college English translation teaching," *International Journal of New Developments in Education*, vol. 2, no. 4, pp. 1-7, 2020.
- [26] P. Zhou and Z. Jiang, "Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin," *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512-524, 2016.

Retraction

Retracted: Rice Drought Damage Assessment Using AMSR-E Data Inversion of Surface Temperature

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

References

- [1] Y. Wang, J. Zhang, and G. Li, "Rice Drought Damage Assessment Using AMSR-E Data Inversion of Surface Temperature," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4117481, 8 pages, 2022.

Research Article

Rice Drought Damage Assessment Using AMSR-E Data Inversion of Surface Temperature

Ying Wang,^{1,2} Jianping Zhang ,^{1,2} and Guochun Li³

¹Chongqing Institute of Meteorological Sciences, Chongqing 401147, China

²Jiangjin Modern Agrometeorological Experiment Station of Chongqing, Jiangjin, Chongqing 402260, China

³Jiangjin Meteorological Administration, Jiangjin, Chongqing 402260, China

Correspondence should be addressed to Jianping Zhang; 2016120257@jou.edu.cn

Received 29 March 2022; Accepted 22 April 2022; Published 17 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Ying 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 extreme drought events caused by global warming have become one of the major issues of general concern all over the world. It is estimated that over the past 50 years, the average annual drought-affected area has reached more than 200,000 km², resulting in a global economic loss of US\$6-8 billion, far exceeding other meteorological disasters. Therefore, conducting real-time and effective drought monitoring research is of great significance for issues such as climate change, drought defense, water resources management, and protection in various regions. Rice is the largest food crop in China and plays a pivotal role in food production. Drought is often regarded as one of the most important stress factors. Scientific, accurate, and timely assessment of the impact of drought on rice yield is essential for improving crop drought resistance and ensuring food production. In this study, based on the meteorological data, rice growth period and yield of the main rice planting areas in Chongqing Yangtze River Basin, and based on the drought index of passive microwave remote sensing observation data (AMSR-E), a statistical model of rice meteorological yield and drought index under the influence of drought is established. A rice drought disaster assessment is carried out. The results of the disaster assessment indicate that under the influence of drought, the rice yield reduction rate of representative sites in Chongqing Yangtze River Basin is between 3% and 10%.

1. Introduction

As a kind of agricultural disaster, dry early morning has a significant impact on the society and economy, especially for rice production. Rice is the main grain type and food crop in China. It is the main food source for the world's Chinese population and is also an important economic source for China's 500 million farmers. Especially in the southern region, which accounts for 40% of China's land area and 60% of the country's population, it provides 90% of the country's rice output. Although the proportion of rice production in the southern region has declined in recent years, the importance of rice in the southern region has not changed [1–3].

Rice production is a kind of agricultural production, which is greatly restricted by natural conditions, especially by precipitation and irrigation conditions. Due to the great differences in geography and climate among different

regions of China, the rice-producing regions in China can be divided into two rice regions, south and north, with the Qinling, Huaihe and Yangtze River lines as the boundaries. The area to the south is the rice area in southern China [4]. The southern region of China is generally rich in precipitation, but because of its vast territory and complex geographical environment, the natural conditions (such as heat and rainfall), and economic and social conditions of the provinces and cities in the southern region are not the same. Production conditions are also different. Among them, drought is one of the important factors affecting rice production in southern China. For example, in 2003, Hunan, Jiangxi, Zhejiang, Fujian, and other provinces in southern China experienced the worst drought since 1971. Since rice is the most water-intensive food crop among crops grown in southern China, it was also affected the most during this drought, and the government and farmers paid a heavy price for it [5–7].

Drought is a water shortage caused by the imbalance of water supply and demand. It has the characteristics of high frequency, long duration, and wide range of influence. Drought index is the basis for studying arid climate, and it is also a key link to measure the degree of drought. At present, there are mainly two types of drought monitoring indicators widely used at home and abroad: one is mainly based on the ground climate data observed at a single point to construct a drought index, and the research on these traditional drought indices is mainly based on deterministic forecasts [8]. It is very easy to be affected by the initial conditions, resulting in inaccurate monitoring results, which will be difficult to effectively reflect the drought situation; the other is the drought monitoring index based on modern satellite remote sensing information, mainly using multi-temporal and multi-spectral, multi-angle remote sensing data through various methods to construct drought index to qualitatively or quantitatively evaluate soil moisture distribution from different aspects. The construction of drought index based on remote sensing information for monitoring has the characteristics of wide range and high spatial resolution, which is of great significance for the monitoring of drought in the human range [9–11].

Soil moisture is an important factor in determining the temporal and spatial dynamic changes of agricultural drought. Soil moisture remote sensing inversion technology overcomes the shortcomings of traditional soil moisture measurement methods and is widely used in agricultural drought monitoring. Both optical remote sensing and microwave remote sensing can successfully detect soil moisture and monitor agricultural drought. Microwave remote sensing is based on the contrast between the dielectric constants of dry soil and water and is considered to be the most suitable monitoring method for soil moisture due to its working characteristics of all-day, all-weather, and certain penetration of soil and vegetation. Active microwave remote sensing has high spatial resolution, but is more susceptible to soil roughness and crops, and is suitable for small-scale soil moisture inversion [11, 12]. Passive microwave remote sensing has a short revisit period and is relatively less affected by roughness and terrain, but the spatial resolution of the images is relatively low, which is suitable for large-scale agricultural drought monitoring.

Compared with infrared and visible light, passive microwave remote sensing has the advantages of long wavelength and strong penetrating ability. Compared with active microwave radar, passive microwave radiometer has the advantages of large monitoring area, short period, less influence by roughness, and less impact on soil moisture. It is more sensitive and the algorithm is more mature. Qiu Yubao et al. used the global surface soil moisture and rainfall rate data retrieved by passive microwave radiometer AMSR-E as the research object, analyzed the impact of precipitation on soil moisture retrieval and its temporal and spatial correlation characteristics, and analyzed the weak correlation between retrieval parameters. The reasons for the emergence were investigated. Alexander et al. studied the influence of soil surface disequilibrium on the use of passive microwave remote sensing to retrieve soil moisture at different scales

(1–4 km) and considered the noise and additional information of the sensor, the brightness of the simulated soil surface conditions [13–15]. The temperature has also been comprehensively studied, and the soil moisture inversion is obtained through the simulation of multi-angle observations by the superposition method of brightness temperature. The uncertainty of soil moisture inversion is mainly caused by the pixel uncertainty of the surface data due to the noise of the sensor. Meanwhile, different spatial scales are studied.

This study takes Chongqing Yangtze River Basin as the research object, uses AMSR-E multi-channel microwave remote sensing data to construct a drought index, establishes a statistical model of rice meteorological yield and drought index under the influence of drought, and conducts rice drought disaster assessment [16, 17].

2. Methods and Theory

2.1. Research Object. Chongqing is located in Southwest China and the upper reaches of the Yangtze River. It crosses the transition zone between the Qinghai Tibet Plateau and the plains in the middle and lower reaches of the Yangtze River between $105^{\circ}11' \sim 110^{\circ}11'$ east longitude and $28^{\circ}10' \sim 32^{\circ}13'$ north latitude. The main rivers in Chongqing include the Yangtze River, Jialing River, Wujiang River, Fujiang River, Qijiang River, Daning River, Apeng River, and Youshui River. The main stream of the Yangtze River crosses the whole territory from west to east, with a flow of 665 kilometers. Chongqing has a subtropical monsoon humid climate, with an annual average temperature of $16 \sim 18^{\circ}\text{C}$, the average temperature in the hottest month of $26 \sim 29^{\circ}\text{C}$ and the average temperature in the coldest month of $4 \sim 8^{\circ}\text{C}$. The annual average precipitation is abundant, with most areas ranging from 1000 mm to 1350 mm. The precipitation is mostly concentrated from May to September, accounting for about 70% of the total precipitation of the whole year. Chongqing has a cultivated land area of 1622000 hectares, with a high degree of agricultural cultivated land development and comprehensive development of agriculture, forestry, animal husbandry, and sideline fisheries. It is an important main grain producing area in China. Among them, the main grain crops include rice, corn, wheat, and sweet potato, especially rice.

2.2. Research Data Sources. The data used by Muwen includes the AMSR-E Level-3 data of each month in Chongqing Yangtze River Basin for 5 years from 2014 to 2019, from the National Snow and Ice Center (NSIDC) in the United States and from all meteorological observation stations in Chongqing Yangtze River Basin from 2010 to 2020. The daily precipitation data and the daily soil moisture data of all soil moisture monitoring stations in Chongqing Yangtze River Basin from 2010 to 2020 are obtained from the China Meteorological Administration.

2.3. AMSR-E Data Presentation. Advanced Microwave Scanning Radiometer AMSR-E (The Advanced Scanning Radiometer for EOS) is an improved design based on the

AMSR sensor and was carried on the Aqua satellite of NASA's Earth Observing System (EOS). AMSR-E is a full-energy passive microwave radiometer with cone scanning. The sensor can observe the range of 89.24N-89.24S. AMSR-E observes the Earth every day from 0 to 8 times, and at the same time every day, there will be some gaps between the swept bands. The AMSR-E data used in this study includes the brightness temperature values of 12 channels, the quality evaluation of the data, and the inversion value of the ground soil moisture. At the same time, the daily data includes orbital ascending and descending orbit data. The main difference between them is that the scanning direction of the sensor is different. The ascending orbit data is scanned from the South Pole to the North Pole, and the descending orbit data is scanned from the North Pole to the South Pole. The spatial resolution of the Level-3 terrestrial product (AELand3) used in this study has been resampled and unified to 25 km.

2.4. Drought Index. When studying the surface vegetation, it is considered that due to photosynthesis, the absorption of light by plant chlorophyll is the reduction of redlight energy reflected by the growth of plants. At the same time, plants absorb little radiation in the infrared band. The reflection is in the infrared. Based on the principle that the energy of the plant increases with the growth of the plant, the vegetation index is constructed by means of the harmonic product quotient of the infrared band and the near-infrared band, and the polarization ratio. Therefore, according to this idea, this study uses the sum-difference quotient and other methods to construct the drought index for the 10 bands under the horizontal and vertical polarization modes of AMSR-E.

The brightness temperature T_b observed by AMSR-E is related to the ground reflectance ε and is expressed as

$$T_b = \varepsilon T_s. \quad (1)$$

Among them, T_b is the surface brightness temperature value received by the satellite, which is related to soil properties, vegetation conditions, and ground roughness, etc., and the influence of soil moisture plays a major role; T_s is the surface temperature. At the same temperature, the microwave radiometer that received surface brightness temperature is related to ε , while ε is closely related to soil moisture content. ε varies from 0.95 to 0.6 from dry soil to water saturated soil. Therefore, the ratio of the information of different bands can be used to try to establish the drought index. The AMSR-E data used in this study lacks data in the 23.8 GHz band, and the selected drought index is TBv06.9/TBv36.5.

2.5. Rice Yield Assessment. The research object of this study is single-cropping rice. The data of rice growth period from agro-meteorological observation stations is collected and analyzed, and the time of transplanting, booting, flowering, and maturity in each rice planting area is obtained by combining with the actual survey data in the field. The rice

transplanting area in Chongqing Yangtze River Basin is mid-June, the booting period is mid-August, the flowering period is early September, and the maturity period is mid-October.

The statistical model of drought index and meteorological yield from transplanting to booting stage, booting to flowering stage, flowering to maturity stage can be expressed as

$$Y_i = A \times D_1 + B \times D_2 + C \times D_3 + D. \quad (2)$$

Y_i is the meteorological yield (kg/hm²); D_1 is the drought index after standardization from transplanting to booting stage; D_2 is the drought index after standardization from booting to flowering stage; D_3 is the drought index after standardization from flowering to maturity; A , B , C is the coefficient; and D is a constant.

The reduction percentage of the actual rice yield after drought and the normal predicted yield without drought is given in the formula

$$Y_d = \frac{(Y_p - Y)}{Y_p}. \quad (3)$$

In the formula: Y_d is the yield reduction rate; Y_p is the normal predicted yield without drought (the standardized drought index for each growth stage in equation (3) is taken as 0); and Y is the historical statistical yield.

3. Results and Discussion

3.1. AMSR-E Drought Index Verification. In this summary, based on the analysis of the actual meteorological conditions in Chongqing Yangtze River Basin from 2014 to 2019, and the distribution characteristics of stations in Chongqing Yangtze River Basin, this study selects a certain station in Chongqing Yangtze River Basin to study the correlation between drought index and drought in time series.

From Figure 1, it can be found that the site drought index is basically positively correlated with the measured soil moisture content. From the standard table of drought grade classification of soil relative humidity (M_s), it can be found that different degrees of drought occurred at site 1 during the test period, and the corresponding drought index reaches a large value, or a maximum value appeared immediately after an upward trend. It can be inferred that when the MPDI index is greater than 0.02, drought may occur in the region.

From the time-series graph of the station, it can be found that the drought index calculated from the AMSR-E brightness temperature data is similar to the change trend of the measured soil moisture, so the index is more suitable for drought monitoring in Chongqing Yangtze River Basin. However, it is predicted time is delayed, and the length of its delay needs to be considered in practical applications.

It can be known from Figure 2 that the site drought index and rainfall basically have a negative correlation trend; the brightness temperature data of the vertical polarization of AMSR-E is larger than that of the horizontal polarization, which is inconsistent with the general theory. Vegetation grows better in the rainy season, which affects the AMSR-E sensor to receive surface radiation, and the vegetation layer

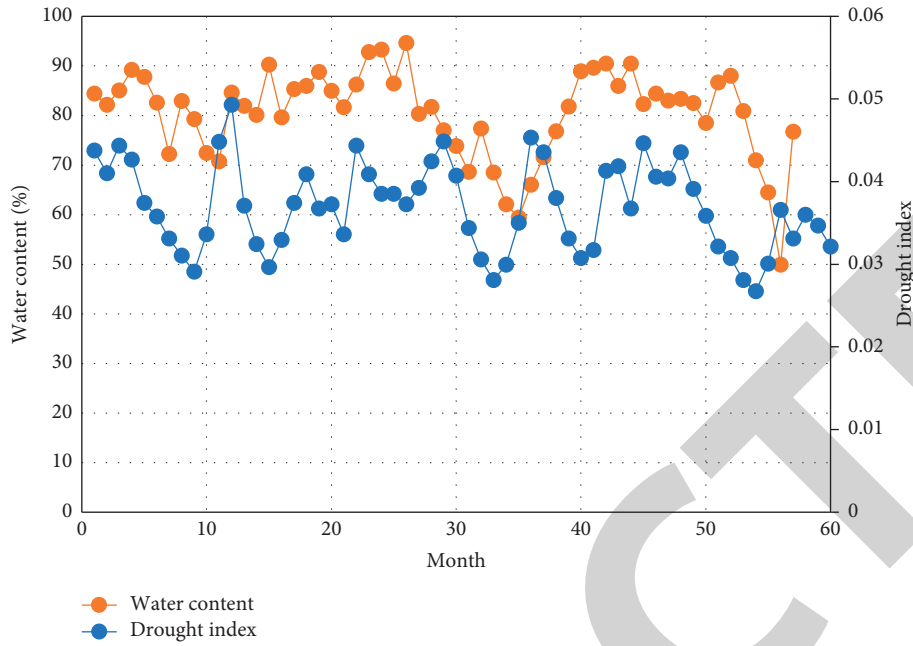


FIGURE 1: Comparison of time series between drought index and measured soil moisture content.

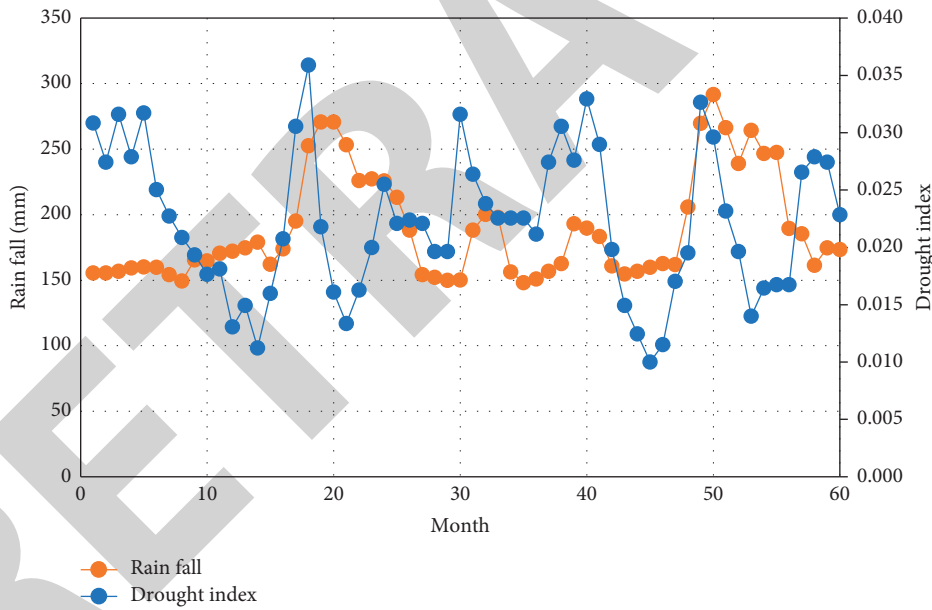


FIGURE 2: Comparison of drought index and rainfall time series.

affects the sensitivity of the brightness temperature measured by the radiometer to the monitoring of soil moisture.

The drought index can reflect the area and degree of drought due to lack of precipitation to a large extent, that is to say, the established ratio drought index can indicate drought to a certain extent. The drought index selected in this study can better reflect the drought caused by the sparse precipitation. Compared with the anomalous vegetation index and conditional vegetation index methods, the microwave method does not need to rely on the surface vegetation and has a high repeated observation rate and is real time. Due to the lower resolution of AMSR-E, the drought index is suitable for larger scale studies.

Due to the difference in solar radiation in each season, the difference in surface temperature causes the difference in surface radiance rate, so the range of dividing the drought index in each season is also different. Another important issue with the ratio drought index is the conversion between the calculated index and true soil moisture. The current research mainly uses the measured soil moisture data for regression, but the regression correlation coefficient is not high and the reliability is poor. Moreover, this method has the problem that the feature space indices of different time and different regions cannot be compared and analyzed. Precipitation and soil moisture are two very important parameters in meteorology, climate, hydrology, and ecological environment, and they are

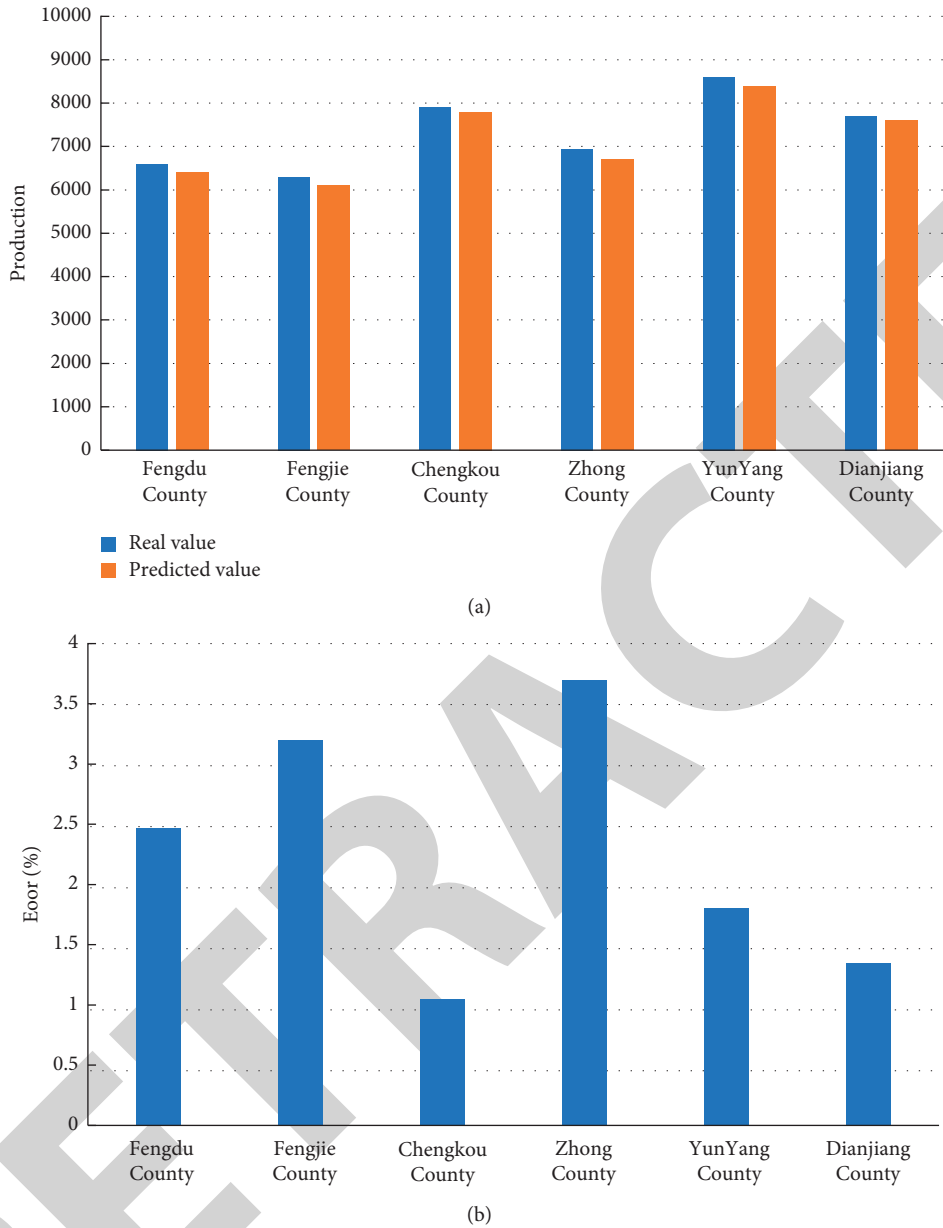


FIGURE 3: Comparison of the product of the real value and predicted value in 2008.

also the most active parts of the global water and energy cycle, and they are also important factors in the occurrence of drought. In the past, traditional drought monitoring methods relying on factors such as precipitation and soil moisture were mainly used, but these factors were few in measurement points, which made it difficult to achieve large-scale, real-time, and dynamic drought monitoring; the modern remote sensing inversion technology currently used is sufficient. Overcoming the shortcomings of traditional soil moisture monitoring methods, it has gradually been widely used in agricultural drought monitoring. Microwave remote sensing is based on the contrast between the dielectric constants of dry soil and water. It has the characteristics of all-day and all-weather monitoring, and has a certain penetrability to soil and vegetation. It is considered to be the most suitable monitoring method for soil moisture. Active microwave remote sensing is

relatively easy to be affected by soil roughness and crops, and the amount of data is large and the data processing is complex, so it is suitable for small-scale inversion; passive microwave remote sensing has a short revisit period and is relatively less affected by roughness and terrain. It has the advantages of large-scale, macroscopic, and dynamic monitoring, but the spatial resolution of its images is relatively low, which is suitable for large-scale drought monitoring.

3.2. Decadal Variation of Drought Index in Each Growth Period

3.2.1. *Transplant to Booting Stage.* From 2014 to 2019, the change trend of drought index from transplanting to booting stage in the six major rice planting areas in Chongqing

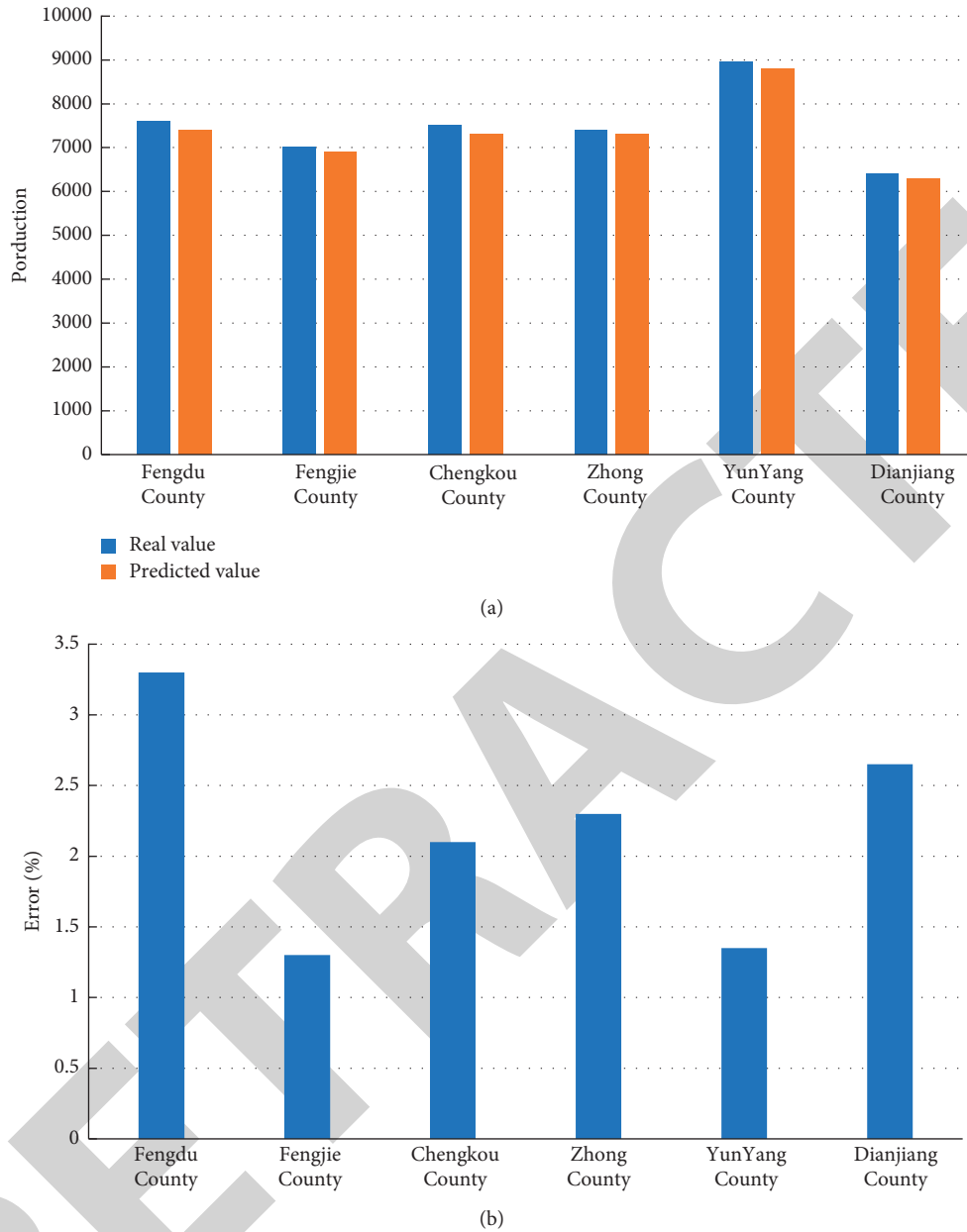


FIGURE 4: Comparison of the product of the real value and predicted value in 2018.

Yangtze River Basin shows that the average drought index of Fengdu County and Dianjiang County is much higher than that of other areas, 1.75 and 0.95, respectively, followed by Zhong County, Chengkou County, Yunyang County, and Fengjie County which are 0.53, 0.45, 0.37, and 0.26, respectively. The drought index showed a fluctuating trend.

3.2.2. Booting to Flowering Stage. The variation trend of drought index from booting to flowering stage in six major rice planting areas in Chongqing Yangtze River Basin from 2014 to 2019 shows that the areas with severe drought from booting to flowering period were mainly Chengkou County, with an average drought index of 2.71; Dianjiang County and Fengdu County followed, with 2.05, 1.73, respectively; Yunyang

County, Fengjie County, and Zhong County had the lowest, with 0.69, 0.55, and 0.41, respectively.

3.2.3. Flowering to Maturity. From 2014 to 2019, the variation trend of drought index from flowering to maturity in the six major rice planting areas in Chongqing Yangtze River Basin shows that the areas with severe drought from flowering to maturity are mainly Dianjiang County and Chengkou County, with the mean values of the drought index being 1.05 and 0.78, respectively, followed by Fengdu County, Fengjie County, Yunyang County, and Zhong County is the lowest.

The above 6 representative statistical models of rice trend yield are from 2014 to 2019 at the site. The trend yield is

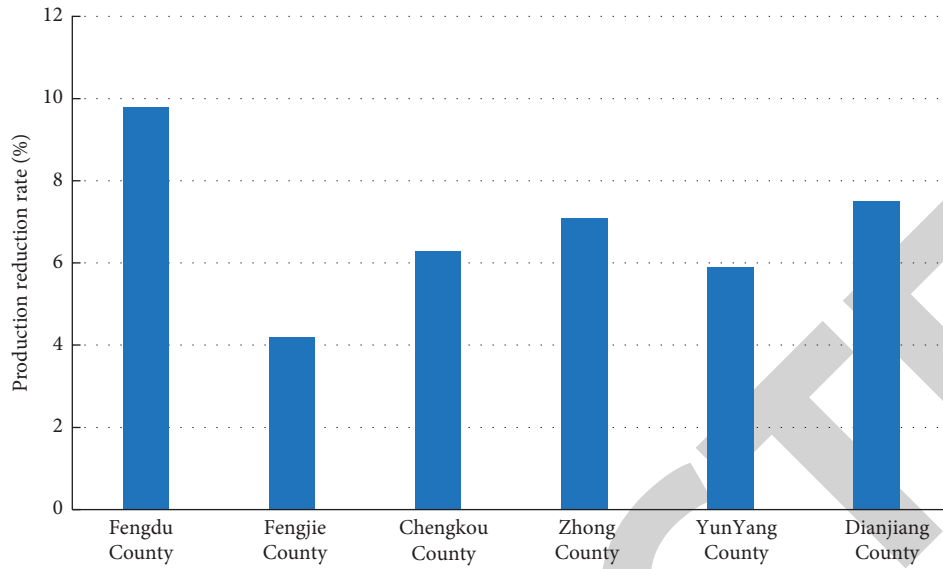


FIGURE 5: Rice disaster assessment in Chongqing Yangtze River Basin in 2018.

obtained by separating the actual yield data using orthogonal polynomials. The statistical models of the rice trend yield of representative sites in different rice planting areas are as follows:

$$\begin{aligned}
 \text{Chengkou County } Y &= -0.05t^3 + 0.689t^2 + 70.234t + 4589.9 \quad R^2 = 0.61, \\
 \text{Fengdu County } Y &= -0.268t^3 + 1.936t^2 + 20.569t + 5196.5 \quad R^2 = 0.66, \\
 \text{Dianjiang County } Y &= 0.459t^3 - 10.265t^2 + 205.673t + 5235.1 \quad R^2 = 0.79, \\
 \text{Zhong County } Y &= -0.698t^3 + 2.615t^2 + 333.216t + 4498.6 \quad R^2 = 0.71, \\
 \text{Yunyang County } Y &= 1.039t^3 - 38.296t^2 + 959.632t + 6103.9 \quad R^2 = 0.79, \\
 \text{Fengjie County } Y &= 0.513t^3 - 21.326t^2 + 323.59t + 5153.6 \quad R^2 = 0.66.
 \end{aligned}
 \tag{4}$$

In drought years, the statistical model established to analyze the relationship between the drought index from transplanting to booting stage, booting to flowering stage, and flowering to maturity stage and meteorological yield is as follows:

$$\begin{aligned}
 \text{Chengkou County } Y &= 70.35 - 365.22D_1 + 202.63D_2 - 99.186D_0, \\
 \text{Fengdu County } Y &= 145.63 - 504.88D_1 + 101.693D_2 - 78.26D_0, \\
 \text{Dianjiang County } Y &= 165.43 - 156.93D_1 + 77.82D_2 - 115.88D_0, \\
 \text{Zhong County } Y &= 169.42 - 226.55D_1 + 89.63D_2 - 102.32D_0, \\
 \text{Yunyang County } Y &= 248.69 - 65.77D_1 + 243.62D_2 - 108.54D_0, \\
 \text{Fengjie County } Y &= 59.52 - 196.52D_1 + 245.36D_2 - 177.81D_0.
 \end{aligned}
 \tag{5}$$

3.3. Rice Drought Disaster Assessment. The comparative analysis represents the difference between the measured rice yield and the simulated yield in the dry years of the site. Figures 3 and 4 show that the absolute value of the relative error between the two is less than 3.5%, which not only shows that the established model can objectively reflect the drought of rice in different growth periods. The impact on

yield can also better assess the loss of rice yield under the influence of drought. At the same time, according to the rice drought disaster assessment model, the measured yield after drought and the normal predicted yield after drought in representative sites in each planting area are analyzed. It can be seen from Figure 5 that under the influence of drought, the rice yield reduction rate of each representative site is 4–10%.

4. Conclusion

- (1) This study analyzes the time-series relationship between the drought index and the daily precipitation index at the measured stations in the study area, and obtains the time-series diagram of the drought index and the daily precipitation and the standardized precipitation index at the measured stations. The drought index selected in this study can fully reflect the drought situation of the site and is suitable for drought research in Chongqing Yangtze River Basin.
- (2) Drought disasters have been one of the main natural disasters in Southwest China since ancient times, and also one of the main agro-meteorological disasters that restrict the sustainable and stable development of agriculture, and have a great impact on the formation of rice yields. In this study, the drought index TBv06.9/TBv36.5 in the MSR-E data is selected based on the comprehensive consideration of crops, atmosphere, and soil in Sichuan rice area, and the drought index is verified by using the actual disaster data, and the effect is good.
- (3) In this study, six representative sites are selected in Chongqing Yangtze River Basin rice planting area for analysis, but the topography in the same area is complex. There are differences in meteorological factors and crop growth periods, and there are certain differences in the evaluation within the

Research Article

Intelligent Classification Method of Archive Data Based on Multigranular Semantics

Xiaobo Jiang 

Jilin University of Architecture and Technology, Changchun, Jilin 130114, China

Correspondence should be addressed to Xiaobo Jiang; jiangxiaobo@jluat.edu.cn

Received 17 March 2022; Revised 9 April 2022; Accepted 15 April 2022; Published 14 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiaobo Jiang. 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 information technology, the amount of data in various digital archives has exploded. How to reasonably mine and analyze archive data and improve the effect of intelligent management of newly included archives has become an urgent problem to be solved. The existing archival data classification method is manual classification oriented to management needs. This manual classification method is inefficient and ignores the inherent content information of the archives. In addition, for the discovery and utilization of archive information, it is necessary to further explore and analyze the correlation between the contents of the archive data. Facing the needs of intelligent archive management, from the perspective of the text content of archive data, further analysis of manually classified archives is carried out. Therefore, this paper proposes an intelligent classification method for archive data based on multigranular semantics. First, it constructs a semantic-label multigranular attention model; that is, the output of the stacked expanded convolutional coding module and the label graph attention module are jointly connected to the multigranular attention Mechanism network, the weighted label output by the multigranularity attention mechanism network is used as the input of the fully connected layer, and the output value of the fully connected layer used to map the predicted label is input into a Sigmoid layer to obtain the predicted probability of each label; then, the model for training: use the multilabel data set to train the constructed semantic-label multigranularity attention model, adjust the parameters until the semantic-label multigranularity attention model converges, and obtain the trained semantic-label multigranularity attention model. Taking the multilabel data set to be classified as input, the semantic-label multigranularity attention model after training outputs the classification result.

1. Introduction

With the development of China's digital archives construction, facing the massive digital archive data, simple statistical methods, or traditional data analysis cannot find the correlation between archive data. Manual classification, compilation, and research of archives also require a lot of manpower and material resources, which takes too long. Therefore, how to discover and use the hidden value of archive data to quickly and accurately classify massive digital archives is a major challenge facing the current archives management field [1–7].

The existing archival data management methods mostly rely on traditional database technology. The goal is to list and organize archival information and basic statistical analysis.

The limitation of database management is that it requires artificial design and analysis content, and the people who formulate the analysis content are required to have a wealth of Experience support. With the rapid update of knowledge, traditional statistical analysis methods cannot meet the association of the content of the excavated archive data and cannot meet the requirements of higher-level intelligent management of archives. At present, natural language processing has become an important branch of artificial intelligence. Natural language processing can classify and cluster digital archives based on the content of digital archives and can well show the changes in the content of digital archives. Among the many association analysis methods, the method based on semantic features provides valuable reference for archive management. Therefore, mining the

content of digital archives is the basis of intelligent archive management. In-depth analysis of the existing digital archive data can better understand the inherent association laws of different types of archives and predict the possible association relationships, so as to provide help for archive managers in archive association analysis and automatic classification [8–12].

Multilabel classification (MLC) is the task of assigning one or more labels to a given input sample. It has a wide range of application scenarios in the real world, such as document annotation, tag recommendation, information retrieval, and dialogue systems. Because the old tags usually have complex dependencies, this makes this task extremely challenging in the field of natural language processing.

Some early research work includes Binary Relevance (BR), Classifier Chains (CC), and Label Powerset (LP) that have achieved good results. As the deep learning based on artificial neural network has made great progress, people have begun to try to use the classic neural network architecture to deal with multilabel classification problems. For example, Zhang et al. used fully connected neural networks and paired ranking loss functions to handle multilabel classification tasks. The overall performance of these methods is to make People satisfied [13–18].

With sequence-to-sequence architecture and attention mechanism, great success has been achieved in neural machine translation tasks. Nam Yang, Lin, and others tried to use them for multilabel classification (MLC) tasks and achieved very advanced results. The sequence-to-sequence model based on attention mechanism uses RNN or long short-term memory network (LSTM) as the encoder to capture the context information in the input sequence, and the decoder also uses RNN or LSTM to generate the label sequence and predict the label. The application of the attention mechanism takes into account the contribution of different parts of the input sequence to the prediction of the label sequence.

The research topic of neural network has received more and more attention in recent years. It is proven to be able to effectively handle tasks with rich relational structures. Shi et al. proposed a multilabel graph convolutional network (ML-GCN) to achieve good results on the task of multilabel node classification. Lanchantin et al. proposed a label information transfer network (LaMP) based on the idea of message transfer in graph neural networks and achieved advanced results on multiple MLC tasks [19–23].

In some early studies, they did not fully consider the impact of the correlation between tags on the prediction of tags or did not fully consider the correlation between tags. For sequence models such as RNN or LSTM, the tag sequence can only be generated sequentially. This makes the sequence-to-sequence architecture limited in time efficiency. In addition, for specific MLC tasks, ideally, sequence factors cannot be considered between output tags. However, the label ranking of the above model is fixed during the training process (usually in descending order), which causes the model to often produce unstable predictions during testing, which reduces the performance and interpretability of the model. For the graph neural network model, only the

relationship between the label and the sample is measured, and the influence of different content within the text on the prediction result is not considered, so it is difficult to apply it to other classic MLC tasks on a large scale [24–27].

Granularity is a tool for describing objects. It is a new concept and computing paradigm for information processing, and the theories, methods, techniques, and tools related to granularity are mainly used to establish a clear model for the intelligent processing of massive amounts of information. For the concept of granules, there is currently no precise definition for granules. From a certain perspective, granules can be regarded as blocks formed by certain relations, such as fuzzy relations, similar relations, and functional relations. People are limited to cognitive ability. When the information to be processed is very complex, the information is divided into information blocks based on logical structure and related relationships, and the information blocks obtained by the division are a granular concept. Within the scope of using the concept of granule, if a granule that cannot be subdivided is defined, it can be called a basic granule. If you use granules to describe files, the smallest unit of information can be called basic granules. As the smallest unit of describing the system, basic granules can be regarded as the most basic constituent elements.

When using the concept of granules, the research object can be divided into a system about granules according to the needs. In the system, there is a certain level of structure, that is, the level of granules, which can be called granularity. Different ways and degrees of granulation produce different levels of granulation systems. Taking a person's archive document as an example, it can be considered that the smallest information is fine-grained, columns are composed of information, tables are composed of columns, and the entire file is composed of tables; conversely, starting from the decomposing of the file, it can be divided into many tables. Granules have the properties of decomposition and assembly. The description of the problem can be carried out from different granularities. Fine granularity means that the granularity is more decomposed, which represents a more detailed description of the problem. On the contrary, coarse-grained means that the decomposition level is small, and it means that there is only a superficial description of the problem. According to the degree of decomposition of the granules, namely, the granularity, the granules can be divided into three levels of granularity: coarse granules, which are the largest parent granules; fine granules, which are basic granules and nondivisible granules; medium granules, which exist between coarse and coarse granules. The grains between the fine grains. Take a file as an example. The entire document is coarse-grained, and the content of a specific column is fine-grained. The part of the information between the coarse-grained and fine-grained can be considered as medium-grained.

In terms of operation and access to the knowledge element, the fine-grained one has advantages, but the operation of the entire document is not enough. The coarse-grained method has a good effect on the operation of the entire document, but for the mining of details, the in-depth knowledge element of processing is not ideal. Medium

granularity is between coarse granularity and fine granularity. If a tree structure is used to describe the document, medium granularity is equivalent to information. In the retrieval system, fragments of this level can be returned. Compared with fine-grained, the medium-grained one saves space and time costs. Compared with coarse-grained, medium-grained method can operate on knowledge elements.

2. Intelligent Classification Method of Archive Data Based on Multigranular Semantics

As shown in Figure 1, this paper introduces a multilabel classification method based on semantic-label multigranularity attention. This article uses this method to apply to the server as an example. It is understandable that the method can also be applied to the terminal. It is applied to include terminals and servers and systems and is realized through the interaction of terminals and servers. The server can be an independent physical server, or a server cluster or distributed system composed of multiple physical servers, or it can provide cloud services, cloud databases, cloud computing, cloud functions, cloud storage, network servers, cloud communications, and intermediate Cloud servers for basic cloud computing services such as software services, domain name services, security services CDN, and big data and artificial intelligence platforms. The terminal can be a smart phone, a tablet computer, a notebook computer, a desktop computer, a smart speaker, a smart watch, etc., but it is not limited to this. The method mentioned in this article includes the following steps:

Semantic-label multigranularity attention model construction: the output of the stacked expanded convolutional coding module and the label graph attention module are connected to the multigranularity attention mechanism network, and the weighted label output by the multigranularity attention mechanism network is used as a full connection. Regarding the input of the layer, the output value obtained by the fully connected layer for mapping the predicted label are input into a Sigmoid layer, and the predicted probability of each label is obtained.

Model training: use the multilabel data set to train the constructed semantic-label multigranularity attention model, adjust the parameters until the semantic-label multigranularity attention model converges, and obtain the trained semantic-label multigranularity attention model; multilabel data set classification: taking the multilabel data set to be classified as input, the semantic-label multigranularity attention model after training outputs the classification result.

In the choice of coding architecture, this paper does not use the traditional CNNRNN or self-attention mechanism but builds the multigranular semantic feature representation of the text sequence hierarchically through the stacked expanded convolution structure. Aiming at the complex correlation between tags, this paper also does not use sequence-to-sequence architecture to generate tag sequences to model but applies a tag graph attention network. It can directly model the label relevance and obtain the updated label representation. In order to efficiently use the semantic

information of multiple levels of word phrases and sentences extracted by the encoder, the present invention also designs a multigranularity attention. It uses these types of hierarchical information to weight the label representations, respectively, highlighting the label representations whose semantic features are closely related to the text sequence. Finally, the weighted label representation is sent to a fully connected layer shared with the weight of the label embedding matrix for linear mapping, and a Sigmoid function is used to realize the output of each label probability.

The specific steps of the method proposed in this article are as follows:

The model built in this article consists of four parts: a stacked dilated convolutional coding module, a label graph attention module that models label correlation, a multigranularity attention mechanism, and the weights of the input label embedding matrix with a Sigmoid function, shared fully connected layer.

2.1. Stacked Dilated Convolution Module. In this paper, the coding architecture is constructed by continuously stacking the expanded convolution structure, which is used to extract the multigranularity semantic feature representation in the input sequence. For a given word embedding matrix representation,

$$\begin{aligned} X &= [x_1, x_2, \dots, x_s], \\ X &= [x_1, x_2, \dots, x_s], \quad X \in R^{S \times d}. \end{aligned} \quad (1)$$

Unlike standard convolution, which operates on a continuous subsequence of the input sequence at each step, expanded convolution has a wider receptive field by adding 8 holes between words within the subsequence, of which 6 is called the expansion rate. Then, the expansion convolution operation for the center word x and the convolution kernel W with width k can be expressed as

$$c_i = \sigma \left(W \bigoplus_{j=0}^k x_{i \pm j\delta} + b \right). \quad (2)$$

where \oplus is the vector splicing and b is the bias term and is the nonlinear activation function. Then, the output of the dilated convolution after the execution of the word vector matrix can be expressed as

$$C = [c_1, c_2, \dots, c_s], \quad (3)$$

where

$$C \in R^{S \times f}. \quad (4)$$

In order to maintain the consistency of input and output, the input sequence is zero-filled before the convolution operation, which effectively preserves the semantic information. As shown in Figure 2, the multigranularity semantic information of the input sequence is extracted by stacking expansion convolution. For the first layer, set $\delta=1$ (equivalent to standard convolution), so that all words in the input sequence will not be missed. Then, the subsequent stacking of the dilated convolution structure uses a larger

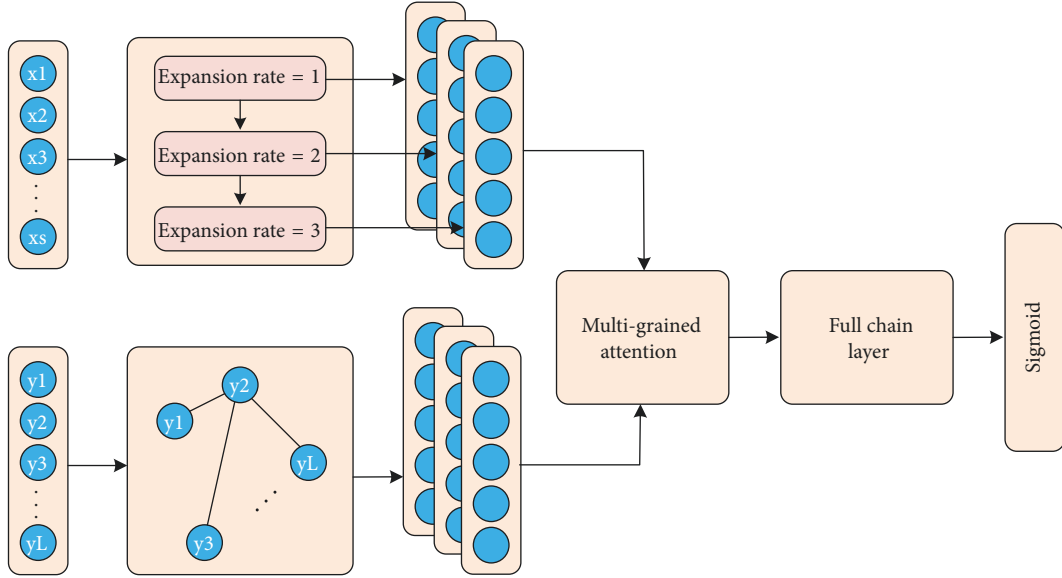


FIGURE 1: Multilabel classification method.

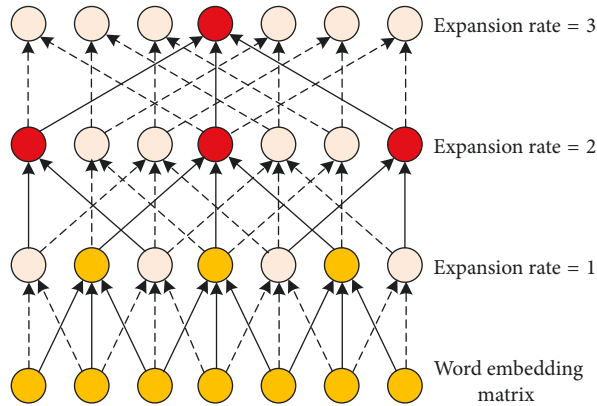


FIGURE 2: The multigranularity semantic information.

dilation rate. The range of the receptive field is expanded in a linear manner, but only a small number of layers and appropriate parameters can cover the entire text sequence.

After each convolutional layer, layer standardization is applied to ensure that the model does not disappear or explode during the training phase. Multilevel expansion rate is designed according to the performance in actual verification. The output of each pair of overlays is a representation of the semantic features of the text at a specific level of granularity. Assuming that the number of stacked layers is L , the output of the first stacked layer is

$$C^l = [c_1^l, c_2^l, \dots, c_s^l], \quad (5)$$

where C is the number of convolution filters in each layer. The multigranularity semantic features of all stacked layers are expressed as

$$[C_1, C_2, \dots, C_L]. \quad (6)$$

In this way, SDC gradually obtains word-level semantic and local semantic features from the word and phrase level

with a small expansion rate and captures the global long-term dependence from the sentence level with a large expansion rate. This coding method not only is superior to traditional sequence models (RNN or LSTM) in parallelism, but also can significantly reduce the amount of parameters and memory loss compared to models based entirely on self-attention mechanisms.

2.2. Label Map Attention Module. This paper uses graph attention networks to model the complex dependencies between tags. A label image is composed of label nodes and adjacency matrix. For a data set with a label space size of L , the adjacency matrix is constructed on the basis of the entire data set, which can be expressed as AER but is different from using statistical knowledge (frequency or word frequency-inverse document frequency) to loosen fixed weights. The parameter label map advocates modeling the correlation between the labels as a weighted map and updates the attention weight and adjacency matrix through the GAT network selective learning method.

The label graph attention network (Label-GAT) takes the label vector obtained through the label embedding matrix WR as input and uses the attention mechanism to participate in the calculation of the weights of neighbor nodes. The calculation process is as follows:

$$h_t^g = \rho \left(\sum_{r \in N_t} \alpha(h_t^{g-1}, h_r^{g-1}) W^{g-1} h_r^{g-1} \right), \quad (7)$$

where α is the attention function, which adaptively controls the contribution of the neighbor node r to the current node t . In order to learn the attention weights of different subspaces, a multihead attention mechanism is used in the Label-GAT network.

$$h_t^g = \parallel_{m=1}^M \rho \left(\sum_{r \in N_t} \alpha_m(h_t^{g-1}, h_r^{g-1}) W_m^{g-1} h_r^{g-1} \right). \quad (8)$$

At the same time, in order to capture the connection between nodes at a greater distance, the Label-GAT network is stacked in multiple layers. The final label representation output through Label-GAT modeling can be expressed as follows:

$$H = [h_1^g, h_2^g, \dots, h_L^g]. \quad (9)$$

2.3. Multigranularity Attention Module. It uses the semantic feature representation C at each granularity to weight the label representation H output by Label-GAT, thereby highlighting the labels that are closely related to the text sequence.

Inspired by the multihead self-attention mechanism, the semantic feature representation under each granularity is regarded as a different subspace, and the number of subspaces (head number) is the same as that of granularity. By calculating the weights of the label representations in multiple subspaces, the weighted results of all granularities are finally spliced, and the weighted label representations are output.

Assume that d represents a Query key-value pair (Key-Value) at the i -th granularity. Then, the output of multi-granularity attention at this granularity is

$$H_{\text{attn}}^l = \text{softmax} \left(\frac{(HW_q^l)(C^l W_k^l)^T}{\sqrt{d_k^l}} \right) (C^l W_v^l). \quad (10)$$

Among them, W is the weight parameter under the first degree. By aggregating the output at all granularities, the final weighted label is expressed as

$$H_{\text{attn}} = [H_{\text{attn}}^1, H_{\text{attn}}^2, \dots, H_{\text{attn}}^n] W. \quad (11)$$

Among them, W is the weight parameter.

2.4. Tag Embedded in the Shared Fully Connected Layer. In order to ensure that the embedding representation of each tag corresponds to the feature tag, a fully connected layer that shares the weight of the tag embedding matrix is

designed. Specifically, after the model uses W to generate the corresponding label vector, it is again used to map the output value of the predicted label. Subsequently, the output value is fed to a Sigmoid layer to generate the predicted probability of each label:

$$\hat{Y} = \text{sigmoid}(W_y H_{\text{attn}}). \quad (12)$$

3. Empirical Analysis

In order to simulate the classification of massive digital archive data, 5 types of text data with a large amount are selected for classification experiment; the categories are C19-Computer, C32-Agriculture, C34-Economy, C38-Politics, C39-Sports, and a total of 6 253 texts of data.

In the experiment process, the preprocessing of word segmentation and stop word removal was first performed on the archive data set, and then the five classification category tags were added to the end of the corresponding archive text as the tags of the training data. Put the file training into FastText for deep learning training.

The specific parameters of FastText used in this article are as follows: set the learning rate to 0.1. Considering the efficiency of text classification, the dimension of the word vector is set to 50, which not only ensures the accurate expression of semantics in the text, but also does not excessively reduce the operating efficiency of the algorithm. On the premise that the data set is sufficient, the FastText algorithm only needs 3 full data sets of training to converge, and it has high training efficiency. Regarding semantic level features $\text{ngarm} = 2$, the classifier uses hierarchical Softmax. The important parameter of the FastText model is the context window (ws). The selection of this parameter means the amount of information that can be obtained from the text context sentence. As the size of the context window increases, the F1 value of text classification will also increase, but training time will increase. For the current data set, when the context window value (ws) is 4, the F1 value of each category of the text will stabilize, reaching about 0.96. In this experiment, FastText model, Naive_Bayes (naive Bayes) model, and SVM model are used for classification training respectively, and the above-mentioned test set is used to verify the classification results. The classification results of the three types of models are shown in Figure 3. According to the evaluation criteria of machine learning classification, the precision rate, recall rate, and F1 value of each of the 5 types of text data are, respectively, obtained, and the macro average, micro average, and weighted average precision rate, recall rate, and recall rate of the comprehensive text test set are calculated. F1 value displays the number of test files of each type at the same time.

The FastText model performed stably in the testing of the test set. In the multiclassification experiment with a total of 6 253 files, the classification accuracy, recall, and F1 value of each category were maintained above 0.94, as shown in Figure 3(a). The performance of the Naive_Bayes model in the test set fluctuates slightly. In the multicategory with a total of 6 253 files, the accuracy, recall and F1 value of each

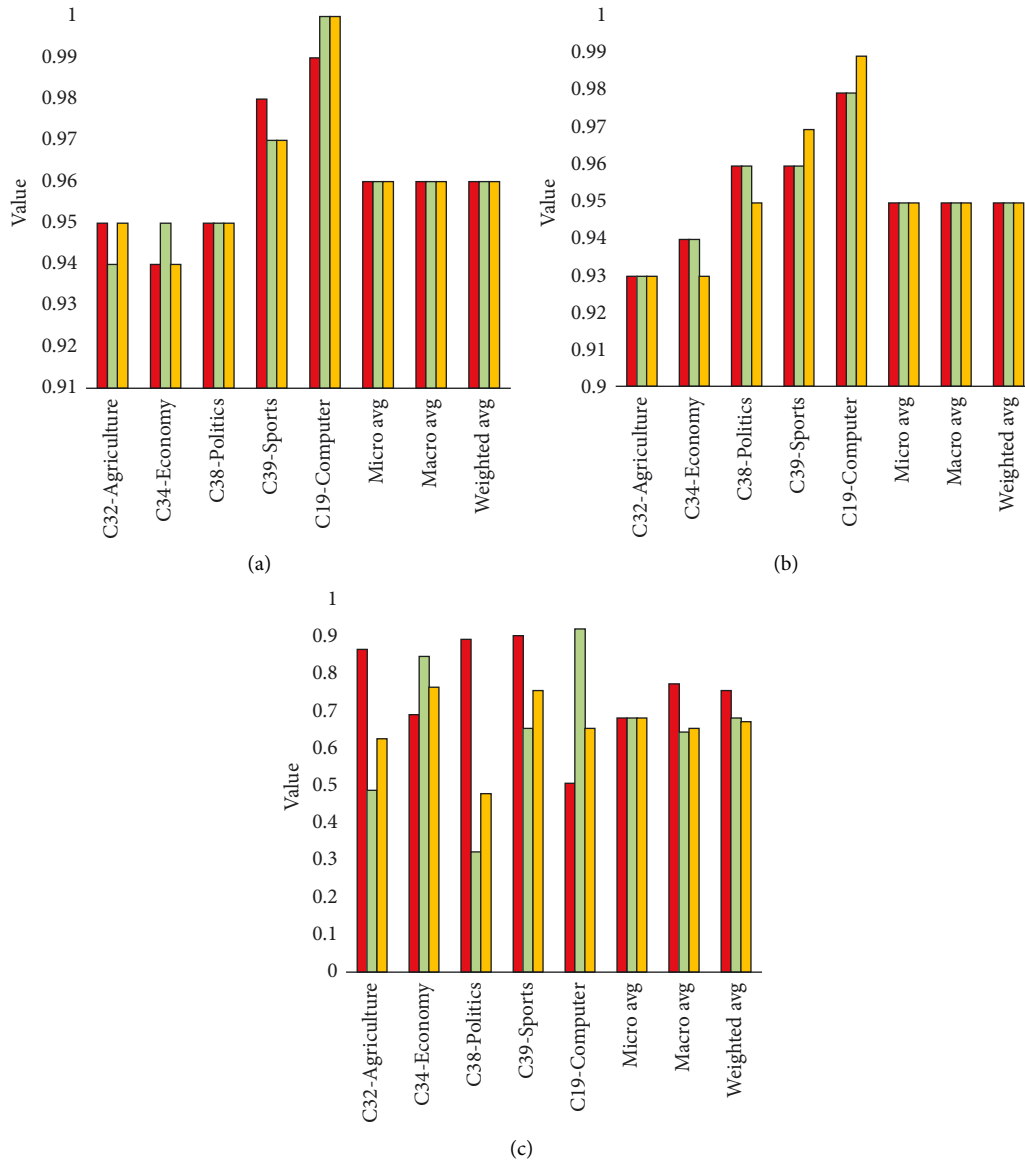


FIGURE 3: Results of classification. (a) FastText, (b) Naive_Bayes, and (c) SVM.

category are maintained above 0.92, as shown in the Figure 3(b). The SVM model performed poorly on the test set. In the multicategory with a total of 6 253 files, the accuracy, recall, and $F1$ value of each category are quite different, and the lowest can only reach 0.52, as shown in Figure 3(c). The predicted also can be found in Figure 4.

Since most of the archive text data is clear, and there is a need for classification and archiving, it is suitable for automatic classification using deep learning. In this experiment, the classification accuracy of the FastText-based text classification model in the three categories of C19-Computer, C32-Agriculture, and C34-Economy is higher than that of the Naive_Bayes model, and the classification of the two categories of C38-Politics and C39-Sports The accuracy is the same as that of the Naive_Bayes model. The reason for this is because the content of the data set used in the experiment is clear, and the original text features are enough to

make the Naive_Bayes model make accurate classification judgments. Even so, the FastText-based classification model in this article can also win in some data categories, which proves the effectiveness of the FastText-based classification model in text classification. The advantage of this classification accuracy will increase with the complexity of the archive data set, which is more prominent.

At the same time, the accuracy of the FastText-based text classification model in all categories is much higher than that of the SVM-based text classification model. From the perspective of the mathematical model, the hyperplane formed by the SVM is not suitable for the text features of the file subject. Although the model is simple, the efficiency of the model is acceptable, but there is still a gap in accuracy compared with the model based on FastText.

It can be seen from Figure 5 that, in the five categories of classification verification experiments, the FastText-based

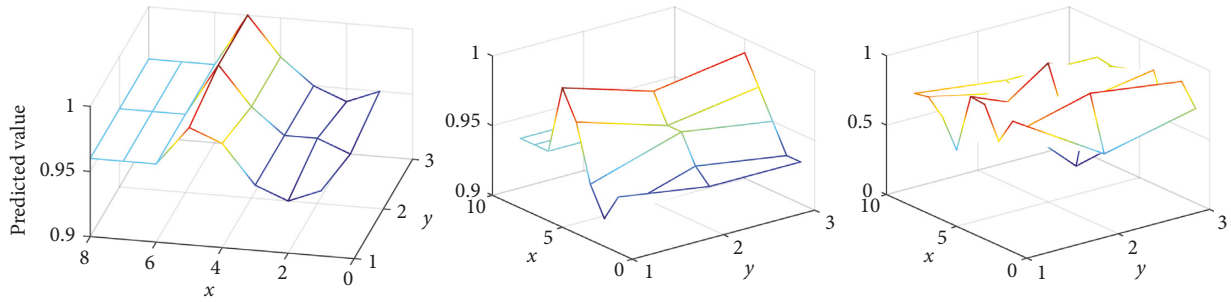


FIGURE 4: Predicted value.

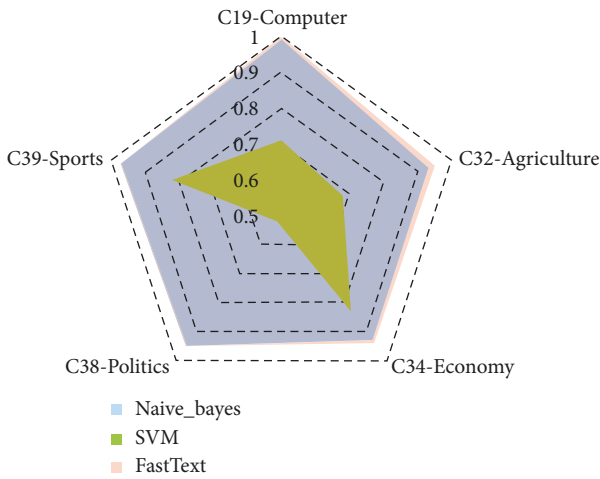


FIGURE 5: Five categories of classification verification experiments.

text classification model has reached the classification F1 values on C19-Computer, C32-Agriculture, C34-Economy, C38-Politics, and C39-Sports as 1.00, 0.95, 0.94, 0.95, and 0.97. It can be seen from the above figure that the weighted F1 value of the naive Bayes classifier is 0.95, and the F1 value of the SVM classifier is 0.73. The text classification model based on FastText is in terms of weighted F1 value, weighted average precision rate, and weighted average recall rate. Among the three overall category evaluation indicators, it is far superior to the text classifier based on the SVM model, which is about 1% improvement over the text classification model based on Naive Bayes, and the overall evaluation F1 value has reached 0.96, basically replacing the manual level of classification.

The label space is fixed to n . Using the transfer method can effectively reduce the label space (from $2n$ to n), reduce the computational complexity, and compress the training time. The experimental strategy is to use single-label data for training and verification and use multilabel data or all data for testing. Therefore, Reuters-21578 was reprocessed, and two sets of experiments were designed. Reuters-Multilabel and Reuters-Full remove all multilabel data in the training set and validation set, leaving only 5912 single-label data. Reuters-Multilabel retains only 437 multilabel data in the test set; Reuters-Full contains all 3019 data.

A comprehensive comparison is made between the proposed model SLMA, capsule network, and SGM. Use

MicroF1 score and MacroF1 score as evaluation indicators instead of Hamming loss. In addition, we have introduced precision (Precision) and recall (Recal1) for a more comprehensive comparison. The results show that, on the test sets of the Reuters-Multilabel and Reuters-Full data sets, our method has significant advantages over Capsule and SGM in all four evaluation indicators. Especially on the Reuters-Multilabel dataset, which only contains multilabel data, a great improvement has been made in the testing process. Different from the capsule network, SLMA and SGM focus on predicting tags through the content of different positions in the text, instead of using document-level prediction. SLMA provides a multigranular semantic feature representation that can take into account local semantics and long-term dependence through a stacked expanded convolution structure. It provides richer and finer-grained text information for label prediction, so it is better than SGM in performance. In addition, the good results when using Reuters-Full data for testing also show that SLMA has a strong advantage over competitors in single-label data. Experiments on the robustness of the model when faced with high and low-frequency data: in the experimental results of the above data set, it is observed that the performance of SLMA on low-frequency tags is not as good as high frequency. The y versus x is shown in Figure 6.

Consistent with the prediction, the ratio of the overall sample size to the internal label data caused the performance difference of the model on different label frequencies. In order to observe the overall performance of SLMA more intuitively, two sets of experiments are set up, using MicroF1 and MacroF1 to measure the performance of high-frequency and low-frequency tags, respectively. In addition, for the convenience of comparison, SGM is also used to evaluate RCV1-V2. For the model performance of high-frequency tags, the 10, 20, 30, 40, 50, and 60 tags with the lowest frequency are eliminated. For low-frequency tags, remove the 10, 20, 30, 40, 50, and 60 tags with the highest frequency.

Figure 7 is the results of SLMA and SG on high-frequency and low-frequency tag data. But in the case of tag frequency changes, the SLMA model still maintains a better performance than the strong baseline SGM, and the gap is gradually widening. In contrast, despite the reduction of high-frequency tags, the performance of the two models has different degrees of decline, but our proposed SLMA is more robust than SGM in classifying low-frequency tags. Compared with Seg2Sea-based SGM, SLMA gets rid of the

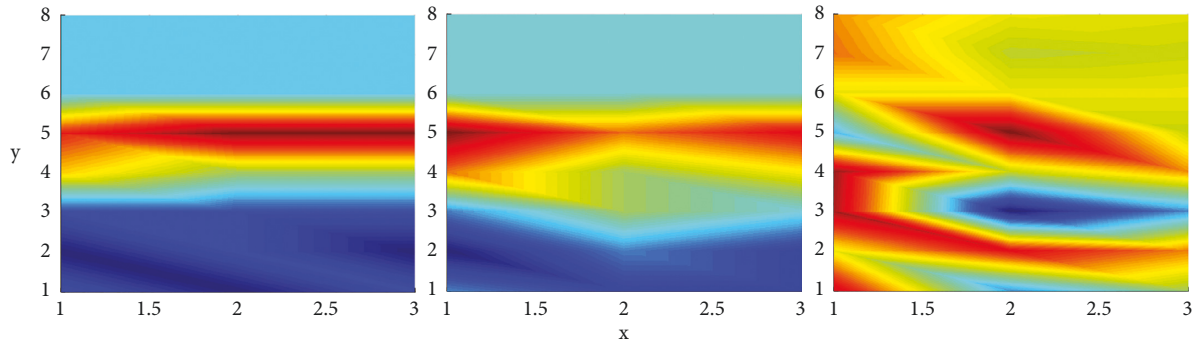
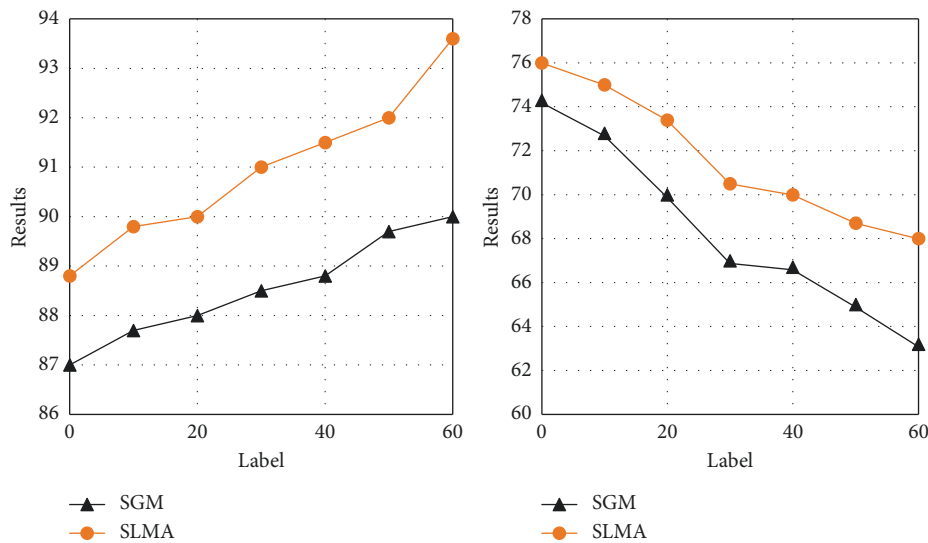
FIGURE 6: y versus x .

FIGURE 7: The results of SLMA and SG. (a) High frequency. (b) Low frequency.

dependence on the order of tag frequency (descending or ascending) and pays more attention to the connection between the multigranular semantic representation of the text and the tags from the perspective of local semantics and long-term dependence. Not only does it perform well in high-frequency tags, but it also more robustly classifies low-frequency tags.

In summary, for multilabel text classification tasks, a semantic-label fine-grained attention model with strong interpretability is proposed. This model constructs a multigranular semantic feature representation of the text sequence in terms of local relevance and long-term dependence through the stacked expanded convolution structure. We conducted a series of experiments on three benchmark data sets, and the results show that the proposed model has very advanced performance in various evaluation indicators. Additional analysis experiments show that our model not only has good migration ability, but also shows strong robustness in the face of high and low-frequency data.

4. Conclusion

This paper proposes an intelligent classification method for archive data based on multigranular semantics. The first is to construct a semantic-label multigranular attention model; that is, the output of the stacked expanded convolutional coding module and the label graph attention module are connected to the multigranular attention mechanism network. The weighted label output by the multigranular attention mechanism network is used as the input of the fully connected layer, and the output value of the fully connected layer used to map the predicted label is input into a Sigmoid layer to obtain the predicted probability of each label; then, the model is trained. Use the multilabel data set to train the constructed semantic-label multigranularity attention model, adjust the parameters until the semantic-label multigranularity attention model converges, and obtain the trained semantic-label multigranularity attention model. Taking the multilabel data set to be classified as input, the semantic-label multigranularity attention model after

training outputs the classification result.

The research basis of this paper is predefined rules. However, the research on predefined rules is not perfect, and a unified definition of rules has not been formed. We can make an in-depth analysis in this aspect in the future.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The author thanks Social Science Research Planning Project of the Education Department of Jilin Province: Analysis on the Digital Management of Private Universities in the 5G Era (no. JJKH20221219SK).

References

- [1] B. Tang, H. He, P. M. Baggenstoss, and S. Kay, "A bayesian classification approach using class-specific features for text categorization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 6, pp. 1602–1606, 2016.
- [2] Y. Liu, L. Li, Y. Tu, and Y. Mei, "Fuzzy TOPSIS-EW method with multi-granularity linguistic assessment information for emergency logistics performance evaluation," *Symmetry*, vol. 12, no. 8, pp. 21–29, 2020.
- [3] W. Lu, X.-P. Wang, J. Zhao, and Y.-K. Zhai, "Research on Teleconsultation service quality based on multi-granularity linguistic information: the perspective of regional doctors," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, p. 113, 2020.
- [4] Y. Zhai and Z. Xu, "Managing individual evaluator's personalized semantic environment of linguistic term with improved vector expression in multi-granularity linguistic group decision making," *Applied Soft Computing*, vol. 92, pp. 106334–106338, 2020.
- [5] F. Meng, D. Zhou, and X. Chen, "An approach to hesitant fuzzy group decision making with multi-granularity linguistic information," *Informatica*, vol. 27, no. 4, pp. 767–798, 2016.
- [6] J. Lin, R. Chen, and Q. Zhang, "Similarity-based approach for group decision making with multi-granularity linguistic information," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 24, no. 6, pp. 28–35, 2016.
- [7] P. D. MacIntyre, T. Gregersen, and S. Mercer, *Positive Psychology in SLA*, Multilingual Matters, Bristol, 2016.
- [8] Z. Wang, H. Ren, Q. Shen, W. Sui, and X. Zhang, "Seismic performance evaluation of a steel tubular bridge pier in a five-span continuous girder bridge system," *Structures*, vol. 31, no. 1, pp. 909–920, 2021.
- [9] Q. Zhao, S. Lyu, Y. Li, Y. Ma, and L. Chen, "MGML: multigranularity multilevel feature ensemble network for remote sensing scene classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 7, pp. 1–15, 2021.
- [10] F. Zhao, Y. Li, L. Bai, Z. Tian, and X. Wang, "Semi-supervised multi-granularity CNNs for text classification: an application in human-car interaction," *IEEE Access*, vol. 2, p. 1, 2020.
- [11] G. Xu, R. Wu, and Q. Wang, "Multi-granular angle description for plant leaf classification and retrieval based on quotient space," *Journal of Information Processing Systems*, vol. 16, no. 3, pp. 663–676, 2020.
- [12] H. Wang, W. H. K. Lam, X. Zhang, and H. Shao, "Sustainable transportation network design with stochastic demands and chance constraints," *International Journal of Sustainable Transportation*, vol. 9, no. 2, pp. 126–144, 2015.
- [13] S.-M. Hosseinasab and S.-N. Shetab-Boushehri, "Integration of selecting and scheduling urban road construction projects as a time-dependent discrete network design problem," *European Journal of Operational Research*, vol. 246, no. 3, pp. 762–771, 2015.
- [14] Y. Dong, W. Yang, J. Wang, J. Zhao, and Y. Qiang, "MLW-gcForest: a multi-weighted gcForest model for cancer subtype classification by methylation data," *Applied Sciences*, vol. 9, no. 17, pp. 1–4, 2019.
- [15] Z. Wang, Y. Zhang, M. Yu et al., "Multi-granular text encoding for self-explaining categorization," *CoRR*, vol. 1, no. 3, 2019.
- [16] C. Zhang, C. Cui, S. Gao et al., "Multi-gram CNN-based self-attention model for relation classification," *IEEE Access*, vol. 7, pp. 5343–5357, 2019.
- [17] X. Liu, R. Wang, Z. Cai, Y. Cai, and X. Yin, "Deep multi-grained cascade forest for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 8169–8183, 2019.
- [18] Q. Guo, Z. Zhu, Q. Lu, D. Zhang, and W. Wu, "A dynamic emotional session generation model based on Seq2Seq and a dictionary-based attention mechanism," *Applied Sciences*, vol. 10, no. 6, p. 1967, 2020.
- [19] H. Ren, X. Mao, W. Ma, J. Wang, and L. Wang, "An English-Chinese machine translation and evaluation method for geographical names," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, pp. 193–201, 2020.
- [20] S. Senthil Kumar and H. Hannah Inbarani, "Cardiac arrhythmia classification using multi-granulation rough set approaches," *International Journal of Machine Learning and Cybernetics*, vol. 9, no. 4, pp. 651–666, 2018.
- [21] M. Molinera Juan Antonio, J. Mezei, C. Carlsson, and E. Herrera-Viedma, "Improving supervised learning classification methods using multigranular linguistic modeling and fuzzy entropy," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 5, pp. 1078–1089, 2017.
- [22] W. Pan, K. She, and P. Wei, "Multi-granulation fuzzy preference relation rough set for ordinal decision system," *Fuzzy Sets and Systems*, vol. 312, pp. 87–108, 2017.
- [23] L. Bote-Curiel, S. Muñoz-Romero, A. Gerrero-Curieses, and J. L. Rojo-Álvarez, "Deep learning and big data in healthcare: a double review for critical beginners," *Applied Sciences*, vol. 9, no. 11, p. 2331, 2019.
- [24] J. Zhang and T. Matsumoto, "Corpus augmentation for neural machine translation with Chinese-Japanese parallel corpora," *Applied Sciences*, vol. 9, no. 10, pp. 1–12, 2019.
- [25] A. T. Azar, S. S. Kumar, H. H. Inbarani, and A. E. Hassaniien, "Pessimistic multi-granulation rough set-based classification for heart valve disease diagnosis," *International Journal of Modeling, Identification and Control*, vol. 26, no. 1, pp. 42–51, 2016.
- [26] Q. Zhang, T. Zhang, and K. Sadarangani, "Binary classification of multigranulation searching algorithm based on probabilistic decision," *Mathematical Problems in Engineering*, vol. 2016, Article ID 9329812, 14 pages, 2016.
- [27] X. Xiao, "Analysis on the employment psychological problems and adjustment of retired athletes in the process of career transformation," *Modern Vocational Education*, vol. 5, no. 12, pp. 216–217, 2018.

Research Article

Analysis and Prediction of Subway Tunnel Surface Subsidence Based on Internet of Things Monitoring and BP Neural Network

Baitian Wang,² Jing Zhang,³ Longhao Zhang,⁴ Shi Yan ,⁵ Qiangqiang Ma ,⁶ Wentao Li,⁶ and Maopeng Jiao⁷

¹*School of Architecture and Civil Engineering, Shangqiu University, Shangqiu 476113, Henan, China*

²*School of Civil Engineering, Nanjing Tech University, Nanjing 210000, Jiangsu, China*

³*Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA*

⁴*Huludao Highway Survey and Design Institute Co. Ltd., Huludao 125003, China*

⁵*Modern Educational Technology Center, Mudanjiang Medical University, Mudanjiang, Heilongjiang 157011, China*

⁶*School of Civil Engineering, Qingdao University of Technology, Qingdao 266033, China*

⁷*School of Civil Engineering, Southeast University, Nanjing 211189, China*

Correspondence should be addressed to Shi Yan; yanshi@mdjmu.edu.cn

Received 7 March 2022; Revised 21 March 2022; Accepted 31 March 2022; Published 14 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Baitian 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 acceleration of the urban development process and the rapid growth of China's population, the subway has become the first choice for people to travel, and the urban underground space has been continuously improved. The subway construction has become the focus of urban underground space development in the 21st century. During the construction of subway tunnels, the problem of surface settlement will inevitably be caused, and the problem of surface settlement will have a certain safety impact on the safe use of surface buildings. The impact of surface construction is predicted, so as to select the best construction technology and avoid the problem of surface subsidence to the greatest extent. On the basis of analyzing the principle of surface subsidence, this paper studies the optimal control strategy and process of subsidence in subway tunnel engineering. The research results of the article show the following. (1) The two sections of the pebble soil layer have basically the same subsidence trend. Among them, the first section has a larger settlement amplitude and both sides are steeper. The second section is mainly cobble clay soil. The pebble layer has good mechanical properties. If it can be well filled, its stability will be improved. The comparative analysis of the two sections shows that with the increase of the soil cover thickness, the maximum subsidence at the surface gradually decreases. The reason is that when the stratum loss is the same, the greater the soil cover thickness, the greater the settlement width. Sections 2 and 3 of a single silty clay have relatively close settlement laws, and the settlement changes on both sides of the tunnel are similar. (2) The surface subsidence caused by the excavation of the side hole accounts for more than 50% of the total surface subsidence, and the width of the settlement tank after the excavation of the side hole is increased by 8–10 meters compared with the excavation of the middle hole. (3) The prediction error of the BP neural network model proposed in this paper is the lowest among the four models, whether it is the prediction of the cumulative maximum surface subsidence or the location of the cumulative maximum surface subsidence, and the average relative error of the cumulative maximum surface subsidence is 3.27%, the root mean square error is 3.87, the average relative error of the location of the cumulative maximum surface subsidence is 7.96%, and the root mean square error is 21.06. In the prediction process of the cumulative maximum surface subsidence, the prediction error value of the Elman neural network is relatively large, and the GRNN generalized neural network and RBF neural network have no significant changes; in the process of predicting the position where the cumulative maximum surface subsidence occurs, the prediction error value of RBF neural network is maximum.

1. Introduction

With the acceleration of urban development, our country's subway construction has also entered a period of rapid development. In order to meet the convenience of the general public, subway tunnels often pass through the city center. The impact of subway construction on the surface may lead to the safety and use of surface buildings. In order to avoid and eliminate this impact as much as possible, we must predict the impact of subway construction on surface buildings, so as to choose the best construction technology that can avoid the problem of surface subsidence to the greatest extent. There are many factors that affect the surface subsidence. In the process of subway construction, the problem of surface subsidence is of great concern. This article will analyze and study the law of subway subsidence and make predictions of surface subsidence. Hui-Jun and Heng [1] analyzed the deformation law and the influence of construction sequence on the surface subsidence based on the monitoring and measurement of the surface subsidence. Wang et al. [2] took the Wuhan Metro Line 3 project as the background, adopted the method of combining on-site monitoring and numerical simulation, and compared and analyzed the vertical surface subsidence law. Li and Shaw [3] analyzed and summarized the methods of controlling the subsidence of borehole shallow tunnels under the surface structure under the background of engineering design and construction. Hui [4] took the Chehuang Station section of the northern extension of Guangzhou Metro Line 4 as the engineering background and analyzed the surface subsidence caused by the main construction parameters of the shield tunnel. Jie et al. [5] took the whole area parallel to the tunnel as an example and analyzed the influence of five basic calculations. Yu and Province [6] combined the case of surface settlement of Wuhan Line 7 shield tunnel excavation and used mathematical linear analysis and fitting to correct the formula according to its geological conditions and measurement data. Wei [7] took Lanzhou subway tunnel as the engineering background and used GTS-NX software to simulate the construction process. Han [8] analyzed the longitudinal and lateral development and distribution of surface subsidence caused by single and double excavation based on Xi'an Metro Line 1. Xiao-Ping [9] verified the reliability of construction and monitoring in subway construction and provided a reference for subway design and construction. Li et al. [10] took the shield construction of Dalian Metro 202 bid section as an example to monitor and measure the surface settlement and top surface settlement during the shield construction process. Pei et al. [11] used the control variable method to analyze the monitoring data of the construction process and summarized the law of surface settlement. Wang et al. [12], after 3 months of accuracy level measurement, showed that the monitoring results have important guiding and reference significance for the construction of Dalian subway and similar shallow tunnels. Bo [13] applied the deep hole grouting technology to expound the reliability of construction and monitoring technology in subway construction and provided a reference for subway construction. Duan et al. [14] proposed a Gaussian equation

based on the actual measurement in a limited area, which is to predict the soft soil movement caused by shallow tunneling. In [15], in order to study the surface subsidence of Dalian Metro Section 202, a monitoring and measurement station was built during the construction process.

2. Analysis and Prediction of Surface Settlement of Subway Tunnels

2.1. Construction Method of Subway Tunnel. With the continuous improvement and development of subway construction technology, tunnel engineering technology has become more and more abundant, and urban construction methods can generally be divided into three categories. The details are shown in Table 1.

2.2. Factors Affecting Surface Subsidence. According to a large number of engineering practices, there are many factors that affect the surface settlement. Different construction methods, geographical conditions, thickness of covering soil, etc. will affect the surface settlement during tunnel construction. The factors affecting land subsidence are not isolated but interact with each other. The optimal design of construction is inseparable from the professional quality and work level of the construction personnel. In the specific construction process, the staff should check the layers, comprehensively analyze the influence of multiple factors, and take effective engineering measures to reduce the impact of construction as much as possible.

3. Analysis and Prediction of Surface Subsidence

3.1. Analysis of Surface Subsidence. Since tunnel excavation is a plane strain problem, according to the stochastic theory, the final ground subsidence ${}_eW(X)$ and surface horizontal displacement $U_e(X)$ at the distance X from the center of the unit caused by unit excavation are [19]

$$\begin{aligned} W_e(X) &= \frac{\tan \beta}{\eta} \exp \left[-\frac{\pi \tan^2 \beta}{\eta^2} X^2 \right] d\epsilon d\eta, \\ U_e(X) &= \frac{X \tan \beta}{\eta^2} \exp \left[-\frac{\pi \tan^2 \beta}{\eta^2} X^2 \right] d\epsilon d\eta, \end{aligned} \quad (1)$$

where β is the main influence angle of the upper stratum of the tunnel, $O - \epsilon\eta$ is the coordinate system adopted by the excavation unit rock and soil mass, and $O - XYZ$ is the coordinate system adopted by the ground surface.

The surface settlement $W(X)$ and the surface horizontal displacement $U(X)$ caused by the excavation of the tunnel are, respectively [20],

$$\begin{aligned} W(X) &= \int \int_{\Omega-\omega} \frac{\tan \beta}{\eta} \exp \left[-\frac{\pi \tan^2 \beta}{\eta^2} (X - \epsilon)^2 \right] d\epsilon d\eta, \\ U(X) &= \int \int_{\Omega-\omega} \frac{(X - \epsilon) \tan \beta}{\eta^2} \exp \left[-\frac{\pi \tan^2 \beta}{\eta^2} (X - \epsilon)^2 \right] d\epsilon d\eta. \end{aligned} \quad (2)$$

TABLE 1: Comparison of several construction methods.

Comparison indicator	Cut-and-cover method [16]	Shield method [17]	Shallow burial and underground excavation method [18]
Geology	Various strata available	Various strata available	Water-bearing formations require special treatment
Place	Occupies more road area	Takes up less road area	Takes up less road area
Section change	Good adaptability	Poor adaptability	Good adaptability
Buried position	Shallow burial	Need some depth	Shallow burial
Waterproof construction	Easier	Easy	Difficult
Subsidence	Smaller	Smaller	Larger
Traffic obstacle	Greater impact	Less affected	Less affected
Underground pipeline	Demolition and protection required	No demolition and protection required	No demolition and protection required
Construction noise	Big	Small	Small
Surface demolition	Big	Small	Small
Water treatment	Precipitation, drying	Block and drop combination	Block, drop, or block-drain combination
Schedule	Affected by demolition and relocation, and the total construction period is relatively fast	The preliminary project is complex and the total construction period is average	Fast start and slow total construction period

According to the replacement formula of double integral, we can get

$$\iint_D f(\varepsilon, \eta) d\varepsilon d\eta = \iint_{D'} f[X_1 + r \cos \theta, Z - r \sin \theta] r dr d\theta. \quad (3)$$

Subsidence value:

$$W(X) = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} \left\{ \frac{\tan \beta}{Z_1 - r \sin \theta} \cdot \exp [X - (X_1 + r \cos \theta)]^2 \right\} r dr d\theta. \quad (4)$$

Surface level displacement:

$$U(X) = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} \left\{ \frac{[X - (X_1 + r \cos \theta)] \tan \beta}{(Z_1 - r \sin \theta)^2} \cdot \exp \left\{ \frac{\pi \tan^2 \beta}{(Z_1 - r \sin \theta)^2} [X - (X_1 + r \cos \theta)]^2 \right\} \right\} r dr d\theta. \quad (5)$$

Surface subsidence and deformation formula:

$$W(X) = \sum_{i=1}^n W_i(X), \quad (6)$$

$$U(X) = \sum_{i=1}^n U_i(X).$$

Surface subsidence and surface horizontal displacement caused by the construction of the double-hole circular tunnel [21]:

$$W(X) = W_1(X) + W_2(X), \quad (7)$$

$$U(X) = U_1(X) + U_2(X).$$

Horizontal deformation:

$$E(X) = \frac{dU_1(X)}{dX} + \frac{dU_2(X)}{dX}. \quad (8)$$

Tilt:

$$T(X) = \frac{dW_1(X)}{dX} + \frac{dW_2(X)}{dX}. \quad (9)$$

Curvature:

$$K(X) = \frac{d^2W_1(X)}{dX^2} + \frac{d^2W_2(X)}{dX^2}. \quad (10)$$

3.2. Prediction Model Establishment. The mathematical model of the settlement tank of the tunnel construction surface cross section follows the normal distribution curve, and the calculation formula of the settlement tank is

$$s = s_{\max} \exp\left[-\frac{y^2}{2i^2}\right], \quad (11)$$

where y is the distance from the centerline of the tunnel, s is the surface subsidence at y from the centerline of the tunnel, s_{\max} is the maximum subsidence at the centerline of the tunnel, and i is the distance from the inflection point to the centerline of the tunnel.

It can be concluded that

$$\ln\left(\frac{s}{s_{\max}}\right) = \frac{-1}{2i^2}y^2. \quad (12)$$

Settling tank width parameters:

$$i = \sqrt{\frac{1}{2m}}. \quad (13)$$

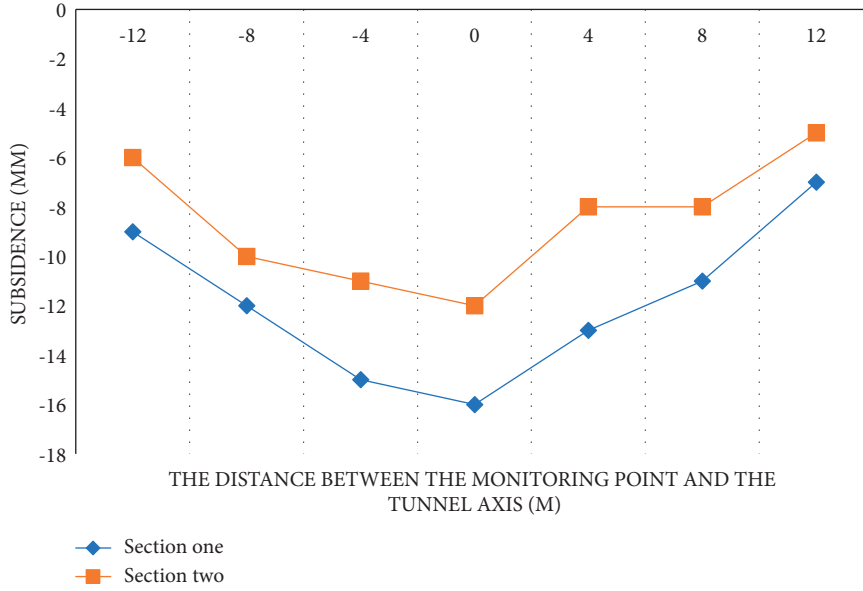


FIGURE 1: Construction settlement curve of pebble soil layer.

The formula fitting of the single-hole settlement tank [22]:

$$s = 16.5 \exp \frac{-y^2}{2 \times 9.53^2}. \quad (14)$$

Shuangdong surface subsidence prediction model [23]:

$$s = s_L + s_R = s_{R \max} \exp \frac{-(y + y_0)^2}{2i_R^2} + s_{L \max} \exp \frac{-(y - y_0)^2}{2i_L^2}. \quad (15)$$

The right hole settlement tank formula fitting:

$$s_R = 16.5 \exp \frac{-(y + 7.5)^2}{2 \times 13.13^2}. \quad (16)$$

The formula fitting of the settlement tank in the left hole:

$$s_L = 27.2 \exp \frac{-(y - 7.5)^2}{2 \times 13.13^2}. \quad (17)$$

The total settling tank formula [24]:

$$s = s_R + s_L = 16.5 \exp \frac{-(y + 7.5)^2}{2 \times 9.53^2} + 27.2 \exp \frac{-(y - 7.5)^2}{2 \times 13.13^2}. \quad (18)$$

4. Simulation Experiments

4.1. Measured Data Analysis. During construction, in addition to having a certain impact on the axis of the excavation surface, it may also have a certain impact on the surrounding environment. In order to increase the reliability of the experiment, the experiment selected different sections of the subway tunnel for research. There are subway tunnels dominated by pebble soil layers and subway tunnels dominated by single silty clay.

4.1.1. Pebble Soil Layer. The subway tunnel in the pebble soil layer is buried about 14 meters deep, and two sections were selected for the experiment. Section 1 is mainly coated with medium and fine sand at the position of shield tunneling, and the upper and lower layers are both pebble layers. Section 2 is buried at a depth of about 15 m. The shield tunneling is mainly composed of pebble layers, and the upper and lower layers are all gravel and silty clay. The construction settlement curve of the pebble soil layer is shown in Figure 1.

According to the data in Figure 1, we can conclude that the two sections have basically the same subsidence trend. Section 1 has a larger subsidence amplitude and both sides are steeper. The main reason is that the stratum where Section 1 is located is mostly fine sand and pebbles. Based on the discontinuous particle distribution, quicksand and other conditions may occur during construction, resulting in a large amount of settlement. This is a typical shallow buried deep section of a shield tunnel, the settlement curve is relatively steep, and there is a large settlement amount. Section 2 is mainly cobble cohesive soil. The pebble layer has better mechanical properties. If it can be well filled, its stability will be better than that of the sand layer. The pebble layer has good water permeability, and the groundwater level in the area where this section is located is low, and the water pressure during construction has little effect on the surface subsidence, so the subsidence curve displayed at the second section of the section has a relatively gentle variation. The comparative analysis of the two sections shows that with the increase of the soil cover thickness, the maximum subsidence at the surface gradually decreases. The reason is that when the stratum loss is the same, the greater the soil cover thickness, the greater the settlement width.

4.1.2. Single Silty Clay. A single silty clay subway tunnel is about 14 m deep [25]. Moreover, the subway tunnel is

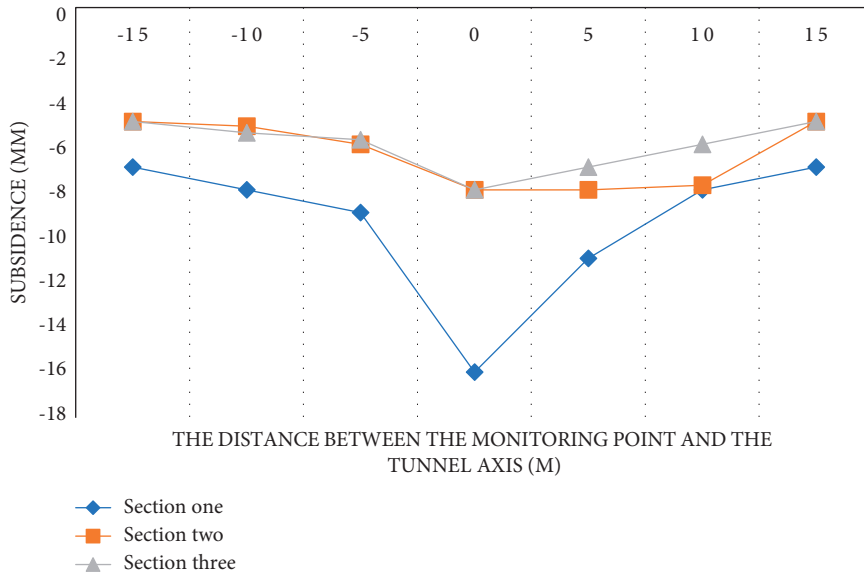


FIGURE 2: Construction settlement curve of single silty clay.

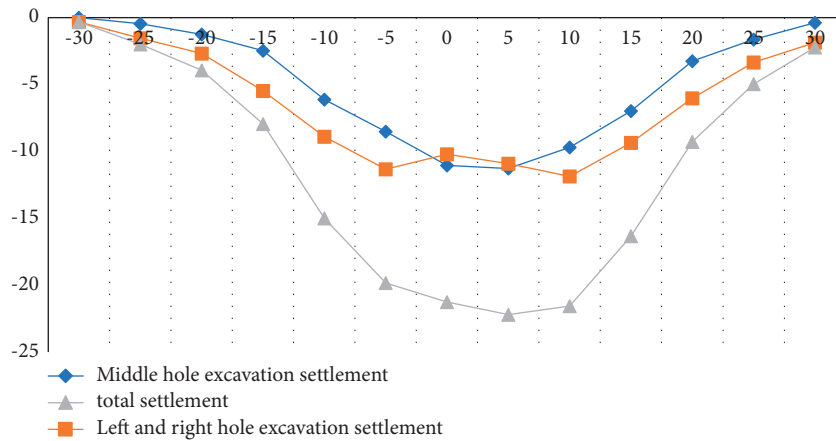


FIGURE 3: Lateral curve of surface subsidence in Section 1.

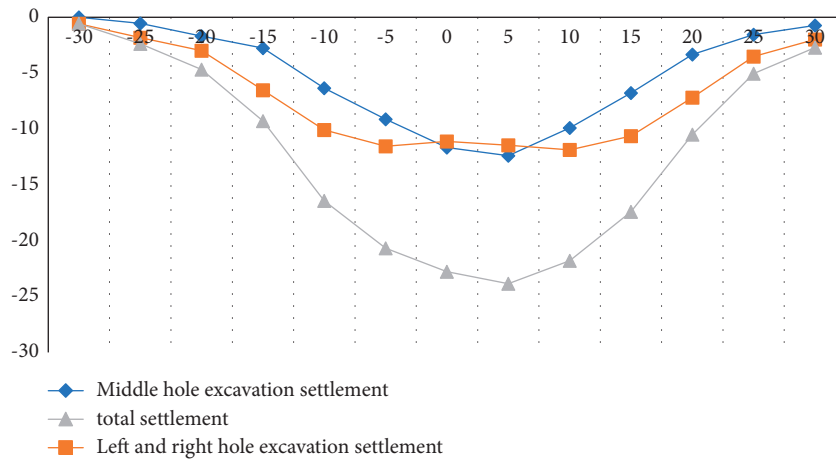


FIGURE 4: Lateral curve of surface subsidence in Section 2.

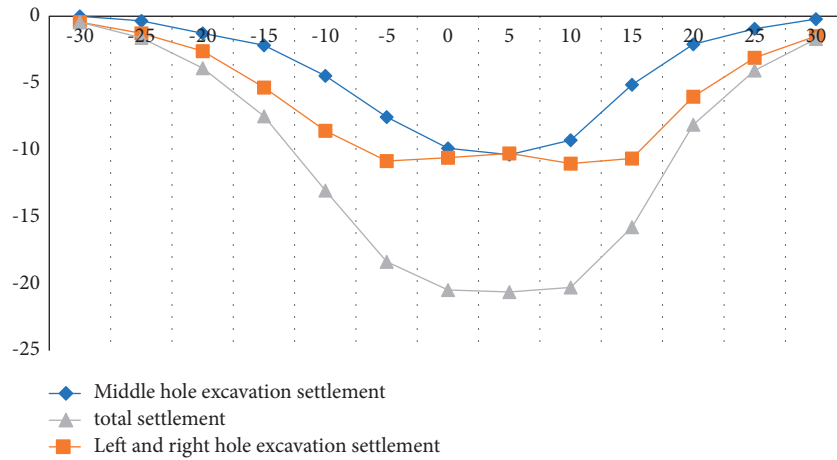


FIGURE 5: Lateral curve of surface subsidence in Section 3.

TABLE 2: Surface subsidence during excavation of each part of Section 1.

Distance from midline	Middle hole excavation settlement	Excavation settlement of left and right side holes	Total settlement	Middle cave settlement/ total settlement	Side tunnel settlement/ total settlement
-25.6	0	-0.32	-0.32	0.0	100.0
-20.6	-0.46	-1.53	-1.99	23.1	76.9
-15.6	-1.25	-2.68	-3.93	31.8	68.2
-10.6	-2.47	-5.47	-7.94	31.1	68.9
-5.6	-6.12	-8.89	-15.01	40.8	59.2
-3.6	-8.51	-11.33	-19.84	42.9	57.1
-1.6	-11.05	-10.21	-21.26	52.0	48.0
1.6	-11.28	-10.92	-22.20	50.8	49.2
3.6	-9.69	-11.87	-21.56	44.9	55.1
5.6	-6.98	-9.35	-16.33	42.7	57.3
10.6	-3.26	-6.02	-9.28	35.1	64.9
15.6	-1.63	-3.34	-4.97	32.8	67.2
20.6	-0.37	-1.86	-2.23	16.6	83.4
25.6	0	-0.49	-0.49	0.0	100.0

TABLE 3: Surface subsidence during excavation of each part of Section 2.

Distance from midline	Middle hole excavation settlement	Excavation settlement of left and right side holes	Total settlement	Middle cave settlement/ total settlement	Side tunnel settlement/ total settlement
-25.7	0	-0.58	-0.58	0.0	100.0
-20.7	-0.55	-1.83	-2.38	23.1	76.9
-15.7	-1.69	-2.98	-4.67	36.2	63.8
-10.7	-2.77	-6.52	-9.29	29.8	70.2
-5.7	-6.34	-10.09	-16.43	38.6	61.4
-3.7	-9.12	-11.55	-20.67	44.1	55.9
-1.7	-11.66	-11.12	-22.78	51.2	48.8
1.7	-12.39	-11.46	-23.85	51.9	48.1
3.7	-9.91	-11.87	-21.78	45.5	54.5
5.7	-6.78	-10.64	-17.42	38.9	61.1
10.7	-3.32	-7.18	-10.5	31.6	68.4
15.7	-1.54	-3.51	-5.05	30.5	69.5
20.7	-0.73	-1.99	-2.72	26.8	73.2
25.7	0	-0.63	-0.63	0	100.0

located in the main city area, so it is greatly affected by external influences during the construction process. A total of 3 typical cross-sections are selected for research in this interval, and the results are shown in Figure 2.

From the experimental data in Figure 2, we can conclude that the second and third sections have relatively close settlement laws, and the settlement changes on both sides of the tunnel are similar. The tunnel is mainly distributed with

TABLE 4: Surface settlement of each part of Section 3 during excavation.

Distance from midline	Middle hole excavation settlement	Excavation settlement of left and right side holes	Total settlement	Middle cave settlement/ total settlement	Side tunnel settlement/ total settlement
-25.8	0	-0.44	-0.44	0.0	100.0
-20.8	-0.34	-1.27	-1.61	21.1	78.9
-15.8	-1.27	-2.61	-3.88	32.7	67.3
-10.8	-2.16	-5.32	-7.48	28.9	71.1
-5.8	-4.46	-8.56	-13.02	34.3	65.7
-3.8	-7.53	-10.82	-18.35	41.0	59.0
-1.8	-9.88	-10.58	-20.46	48.3	51.7
1.8	-10.35	-10.26	-20.61	50.2	49.8
3.8	-9.26	-11.01	-20.27	45.7	54.3
5.8	-5.12	-10.64	-15.76	32.5	67.5
10.8	-2.09	-6.02	-8.11	25.8	74.2
15.8	-0.94	-3.11	-4.05	23.2	76.8
20.8	-0.21	-1.49	-1.7	12.4	87.6
25.8	0	-0.57	-0.57	0	100.0

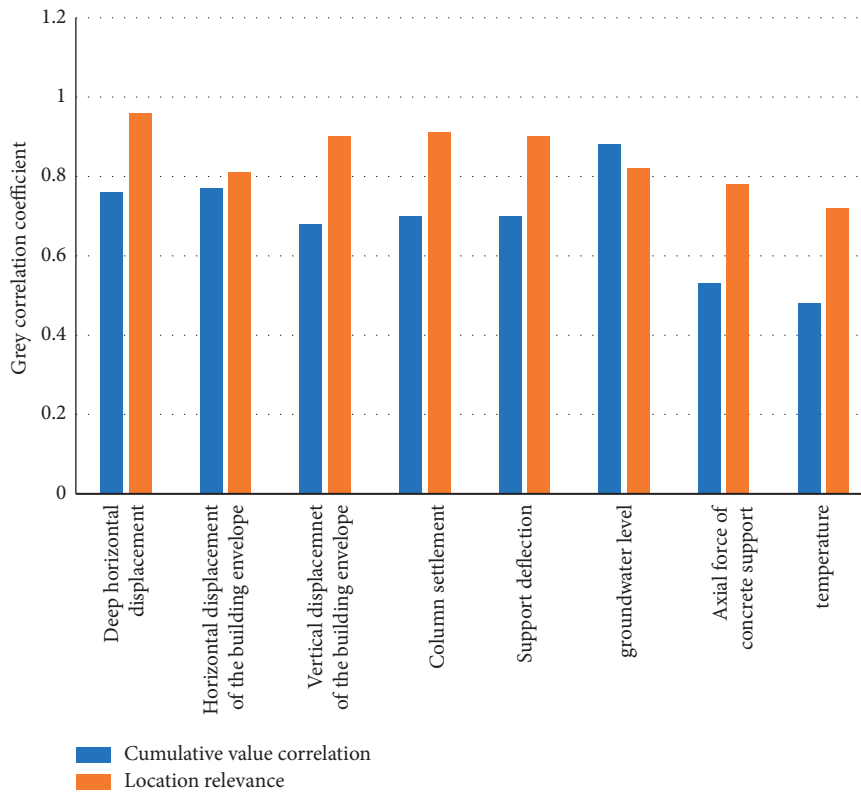


FIGURE 6: Correlation between monitoring content and surface subsidence.

cohesive soil layer within the diameter range, and the geological self-stability is good, so it is not easy to collapse during construction. Due to the better compaction performance, the soil layer is less disturbed during construction in the cohesive soil environment. The settlement value at one section of the section is larger, and the settlement changes on both sides are larger. This section is mainly distributed with thick miscellaneous fill layer, and the stability of surrounding rock is poor. On the whole, Section 1 and Section 2 can better reflect the settlement law of this interval.

To sum up, it can be seen from the above analysis that the mechanical properties of the sand and pebble stratum are

poor, and they are greatly affected during construction and are prone to instability of the excavation surface. When the shield machine is passing through this kind of soil layer, if it encounters pebbles with a large particle size, it is necessary to ensure that the thrust of the shield machine is large enough to push it forward. However, the construction speed will be reduced at this time, and the impact on the surrounding environment will be increased. The sand layer has more coarse sand particles, the thrust increases significantly, and the rotation speed of the cutterhead increases significantly, so the impact on the soil layer during construction is more significant.

TABLE 5: Accuracy and mean square error of each model.

Network type	The maximum accumulated surface subsidence		The location of the cumulative maximum surface subsidence	
	Average relative error (%)	RMS error	Average relative error (%)	RMS error
BP neural network	3.27	3.87	7.96	21.06
GRNN generalized neural network	4.28	10.00	14.38	24.16
RBF neural network	3.27	10.00	17.38	22.70
Elman neural network	7.85	5.29	11.60	18.71

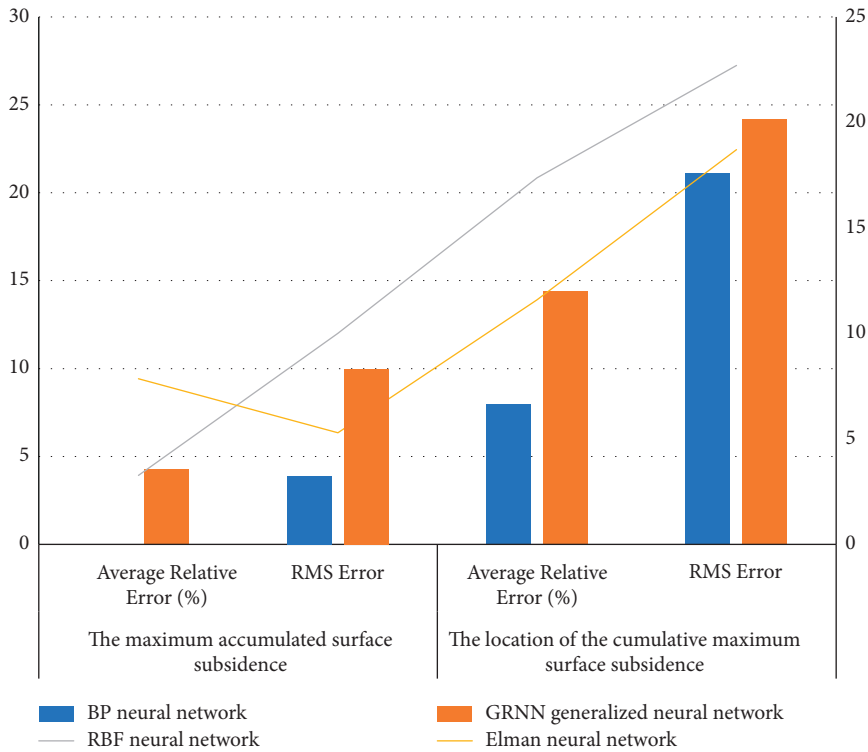


FIGURE 7: Statistics of accuracy and mean square error.

4.2. Analysis of Surface Subsidence. According to a large number of engineering practices and related literatures, there are many factors affecting the surface settlement. According to the research direction of the article, three different sections were selected for the experiment. Surface subsidence is caused by excavation. The experiment records the surface subsidence data of three measurement sections, as shown in Figures 3–5.

According to the above experimental data, we can conclude that during the excavation of the left and right side holes, the surface subsidence curve presents a wave shape, the wave trough is on the center line of the left and right lines, the wave crest is on the central axis of the structure, and the surface subsidence occurs during the excavation of the right side hole. It is larger than the left hole, mainly because when the left hole is excavated first, it will affect the bottom layer on the right side and cause geological softening. The surface subsidence caused by the excavation of the side hole accounts for more than 50% of the total surface subsidence, and the width of the settlement tank after the excavation of the side hole is increased by 8–10 meters

compared with the excavation of the middle hole, as shown in Tables 2–4.

4.3. Model Checking. Different network structures have different performances in predicting surface subsidence. In order to verify the good fitting accuracy and generalization ability of the BP neural network model, the BP neural network proposed in this paper is compared with the GRNN generalized neural network, RBF neural network, and Elman neural network to verify the effectiveness of the BP artificial neural network in surface subsidence prediction. Because the correlation of the experimental detection is not clear, the quantitative relationship between the monitoring content and the surface subsidence should be determined before the formal experiment, as shown in Figure 6.

According to the data in Figure 6, we can conclude that among all the monitoring contents, the correlation between temperature and concrete support axial force and the surface settlement is the smallest, and the other monitoring contents

have strong correlation with the surface settlement. It is almost the same as the position correlation sequence, so it can be seen that the grey correlation analysis can effectively evaluate the correlation between each monitoring content and the surface subsidence. Therefore, in order to avoid adverse effects on the generalization ability of the prediction model, the two monitoring contents of temperature and concrete support axial force are eliminated in this example, and the deep horizontal displacement, the vertical displacement of the enclosure structure, the horizontal displacement of the enclosure structure, and the settlement of the column are used. The supporting deflection, supporting axial force, and groundwater level are used as input variables to predict the surface settlement of foundation pit. We record the relative accuracy value and variance of each model, as shown in Table 5 and Figure 7.

According to the data in Table 5, we can conclude that the prediction error of the BP neural network model proposed in this paper is the lowest among the four models, whether it is the prediction of the cumulative maximum surface subsidence or the prediction of the location of the cumulative maximum surface subsidence. The average relative error of the cumulative maximum value is 3.27%, and the root mean square error is 3.87. The average relative error of the location of the cumulative maximum surface subsidence is 7.96%, and the root mean square error is 21.06. In the prediction process of the cumulative maximum surface subsidence, the prediction error value of the Elman neural network is relatively large, and the GRNN generalized neural network and RBF neural network have no significant changes; in the process of predicting the position where the cumulative maximum surface subsidence occurs, the prediction error value of RBF neural network is maximum. On the whole, the BP neural network model has the best training accuracy and generalization ability for the cumulative maximum surface subsidence and location, and its performance is better than the other three models. It shows that the initial weight and threshold search of BP neural network by artificial bee colony can prevent the network from falling into the local optimal value and effectively improve the network prediction ability.

5. Conclusion

In order to meet the travel problems of the general public and improve the utilization of urban underground space, there are more and more problems in tunnel construction. In the process of urban subway construction, it will inevitably have a certain impact on the rock and soil mass, and the geological deformation may affect the safety of building facilities on the ground. Therefore, the problem of surface subsidence caused by tunnel construction has always been a concern of many researchers. In order to effectively prevent and reduce the problem of surface subsidence caused by subway construction, it is necessary to make accurate predictions of the law of surface subsidence. The BP neural network prediction model proposed in this paper can accurately predict the subsidence law of the surface, and the prediction accuracy is as high as 90%. Since the research on

the law of surface subsidence in this paper is based on the real side of surface data, the data obtained are not sufficient and have certain regional limitations. Therefore, in the future research work, data should be widely collected, and various methods should be used to predict and analyze the surface subsidence.

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 study was supported by the 14th Five-Year Plan of Educational Science in Henan Province (2021yb0474) and the Henan New Engineering Research and Practice Project (2020jglx091).

References

- [1] Y. Hui-Jun and K. Heng, "Analysis of the surface subsidence during the excavating of large-span subway station under shallow cover," *Journal of Railway Engineering Society*, vol. 10, no. 8, pp. 25–30, 2015.
- [2] P. Wang, C. Zhou, Y. Liu, and D. Chen, "A study of land subsidence in the subway shield tunnel excavation in soil strata in Wuhan," *Hydrogeology & Engineering Geology*, vol. 4, no. 13, pp. 69–75, 2015.
- [3] Y. J. Li and H.-M. Shaw, "Pregnenolone and dexamethasone, modulators of cytochrome P450-3A, not increase but reduce urinary α -CEHC excretion in rats," *BioFactors*, vol. 31, no. 1, pp. 67–76, 2007.
- [4] L. Hui, "The shield tunneling parameter optimization and surface subsidence control in che-huang tunnel station sector of Guangzhou subway line 4," *Guangdong Architecture Civil Engineering*, vol. 08, no. 7, pp. 45–56, 2010.
- [5] Y. Jie, Z. Liu, W. Xun, and Y. Li, "Induced surface settlement of stochastic medium theory based on optimization of subway tunnel construction," *Transport Engineering*, vol. 10, no. 12, pp. 114–121, 2016.
- [6] J. Yu and A. Province, "Analysis of surface subsidence around subway in wuhan," *Sichuan Building Materials*, vol. 10, no. 12, pp. 45–59, 2018.
- [7] Z. Wei, "Study on influence of excavation surface distance on surface subsidence of parallel tunnel in subway," *Shanxi Architecture*, vol. 23, no. 9, pp. 45–56, 2017.
- [8] R. Han, "Analysis and prediction for ground subsidence of subway tunnel built in saturated soft loess," *Urban Rapid Rail Transit*, vol. 10, no. 16, pp. 11–21, 2012.
- [9] P. U. Xiao-Ping, "Research on the ground surface subsidence control in the ground fissures segment of subway tunnel," *Construction Quality*, vol. 10, no. 8, pp. 1–7, 2013.
- [10] J. K. Li, B. W. Yan, and Y. J. Shi, "The monitoring and analysis of surface subsidence of soft soil rock large section of subway tunnel shield construction," *Advanced Materials Research*, vol. 848, pp. 78–82, 2013.

- [11] C. Pei, L. I. Juncai, G. Liang, and Z. Jiao, "Multivariate analysis of subway tunnel subsidence regularity," *Journal of Nanjing Tech University (Natural Science Edition)*, vol. 04, no. 12, pp. 21–35, 2014.
- [12] F. F. Wang, H. R. Bai, and J. J. Li, "Analysis of surface subsidence rule of the shallow subway tunnel construction," *Advanced Materials Research*, vol. 671-674, pp. 1081–1086, 2013.
- [13] W. Bo, "Research on surface subsidence control technology of subway tunnel construction at ground fracture section," *IEEE Computer Society*, vol. 10, no. 8, pp. 21–35, 2015.
- [14] S. Duan, L. Huang, Z. Bao, and P. Shen, "Application of modified Peck formula in surface subsidence prediction of Changsha subway tunnel construction," *Journal of Natural Disasters*, vol. 24, no. 1, pp. 164–169, 2015.
- [15] L. I. Jin-Kui, F. F. Wang, and H. R. Bai, "Analysis of surface subsidence rule of shallow subway tunnel construction," *Journal of Wuhan Institute of Technology*, vol. 13, no. 2, pp. 14–21, 2012.
- [16] C. Wan, "Characteristics of strata subsidence and subsidence reduction technology during shield construction of pebble stratum," *Sichuan Architecture*, vol. 12, no. 1, pp. 123–125, 2021.
- [17] G. Chen, L. Wang, and M. M. Kamruzzaman, "Spectral classification of ecological spatial polarization SAR image based on target decomposition algorithm and machine learning," *Neural Computing & Applications*, vol. 32, no. 10, pp. 5449–5460, 2020.
- [18] X. Zhou, "Shield construction technology for subway complex geological conditions," *Engineering Technology Research*, vol. 14, no. 1, pp. 74–75, 2021.
- [19] J. Yu and H. Chen, "Artificial neural network's application in intelligent prediction of surface settlement induced by foundation pit excavation," in *Proceedings of the 2009 Second International Conference on Intelligent Computation Technology and Automation*, pp. 303–305, Changsha, China, October 2009.
- [20] H. Che and J. Wang, "A two-timescale duplex neurodynamic approach to mixed-integer optimization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 36–48, 2021.
- [21] R.-P. Chen, P. Zhang, X. Kang, Z.-Q. Zhong, Y. Liu, and H.-N. Wu, "Prediction of maximum surface settlement caused by earth pressure balance (EPB) shield tunneling with ANN methods," *Soils and Foundations*, vol. 59, no. 2, pp. 284–295, 2019.
- [22] X. Huang, S. Liu, and X. Tan, "An Improved test method for grey relational order," *Journal of Grey System*, vol. 29, no. 1, pp. 64–77, 2017.
- [23] G. Sun, C.-C. Chen, and S. Bin, "Study of cascading failure in multisubnet composite complex networks," *Symmetry*, vol. 13, no. 3, p. 523, 2021.
- [24] 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, 2021.
- [25] K.-H. Park, "Analytical solution for tunnelling-induced ground movement in clays," *Tunnelling and Underground Space Technology*, vol. 20, no. 3, pp. 249–261, 2005.

Research Article

Automatic Recognition Method of Machine English Translation Errors Based on Multisignal Feature Fusion

Ruisi Zhang ¹ and Haibo Huang²

¹*School of International Studies, Hunan Institute of Technology, Hengyang, Hunan 421002, China*

²*School of Electrical Information Engineering, Hunan Institute of Technology, Hengyang, Hunan 421002, China*

Correspondence should be addressed to Ruisi Zhang; zhangruisi@hnit.edu.cn

Received 28 February 2022; Revised 21 March 2022; Accepted 24 March 2022; Published 12 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Ruisi Zhang and Haibo Huang. 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 current automatic recognition method of machine English translation errors has poor semantic analysis ability, resulting in low accuracy of recognition results. Therefore, this paper designs an automatic recognition method for machine English translation errors based on multifeature fusion. Manually classify and summarize the real error sentence pairs, falsify a large amount of data by means of data enhancement, enhance the effect and robustness of the machine translation error detection model, and add the source text to translation length ratio information and the translation language model PPL into the model input. The score feature information can further improve the classification accuracy of the error detection model. Based on this error detection scheme, the detection results can be used for subsequent error correction and can also be used for error prompts to provide translation user experience; it can also be used for evaluation indicators of machine translation effects. The experimental results show that the word posterior probability features calculated by different methods have a significant impact on the classification error rate, and adding source word features based on the combination of word posterior probability and linguistic features can significantly reduce the classification error rate, to improve the translation error detection ability.

1. Introduction

Machine translation refers to a technology that, given a language, uses a certain model or algorithm to translate it into another language that maintains the same semantic information. As a key task in the field of natural language processing, machine translation has played an important role in the communication between countries and nations in recent years. Machine translation is the study of the automatic conversion of one natural language (source language) to another natural language (target language) by means of computers. Machine translation first originated from the idea of using cryptography to solve the task of human language translation proposed by American scientist Weaver in 1949. In 1954, IBM and Georgetown University used electronic computers to translate a few simple Russian sentences into English for the first time. The translation system contains 6 translation rules and 250 words. This experiment demonstrated the process of machine

translation using a dictionary and translation rule-based method. Although it was only a preliminary success, it has since aroused the research fever of machine translation in research institutions in the Soviet Union, the United Kingdom, and Japan, and greatly promoted the research progress of early machine translation. However, in a report entitled LANGUAGE AND MACHINES (Language and Machines) published by the American Automatic Language Processing Advisory Council in 1966, machine translation was completely rejected, which led to the subsequent whole industry and academia began to avoid machine translation. Translation studies suffer from this. Historically, many artificial intelligence research, including machine translation, suffered certain setbacks in that era. The main reason was that the technical level at that time was still relatively low, and the people had high expectations for machine translation and other technologies. Until the mid-to-late 1970s, the exchanges between countries became increasingly close, the communication barriers formed between different

languages became more and more serious, and people's demand for machine translation became more and more intense. At the same time, the development of corpus linguistics and computer science became machine translation that offers new possibilities. Since then, machine translation has entered a period of rapid development. After decades of evolution, it has formed three stages: rule-based machine translation, model-based statistical machine translation, and deep learning-based neural machine translation [1–8]. The transformer framework based on attention mechanism is very effective in translation tasks.

With the more and more frequent exchanges among the various ethnic groups in the world, there is an increasingly urgent need for machine translation to solve the problem of language barriers between various ethnic groups. Machine translation is a project with huge economic and social benefits. Machine translation research requires solving a series of fundamental linguistic problems. The ideal goal of machine translation research is to be practical in a wide range of fields with a high rate of accurate automatic translation. The main obstacle to realizing this ideal goal is the lack of precise natural linguistics systems, description methods, and description tools. To study various linguistic phenomena and laws in natural language, it is particularly necessary to describe language knowledge correctly. The study of natural language is not only important for automatic translation (written, voice, and computer) between natural languages but also for information processing related to the understanding of natural language knowledge levels: automatic summarization, text shaping, automatic formation of practical office letters, information filtering, information selection, network integration services, and other fields are also very important [9–15].

It has been recognized that the acquisition of high-quality machine translation results requires analysis and understanding of the semantics of natural language, and semantic analysis and syntactic analysis should not be disconnected. Therefore, establishing and perfecting the grammar theory suitable for natural language analysis and generation is still a subject of exploration [16]. Some of the existing MT theories involve lexical analysis, some involve syntactic analysis, and some involve semantic analysis. Each has different starting points, different emphases, and each has its own characteristics, and each project has its own limitations. The research method of this paper is mainly based on semantics, based on semantic language theory, combined with syntactic and grammatical analysis, and inherits and develops the existing MT theory.

At present, the research of language technology lags behind computer technology. We should make more efforts in the formal description of language, in terms of semantic description, and establish a rich and complete description of the concept represented by the word, revealing the concept of the concept. A commonsense knowledge base of relationships between concepts and attributes. In the study of English which is easily misunderstood and mistranslated, this paper explores the formal description of language from the perspective of unified semantics (rather than from the perspective of a natural language) [17–20]. The syntactic-

semantic tree and syntactic-semantic type tree proposed in this paper are concept hierarchy, which can reveal the syntactic and semantic relationship between concepts. The computer's English translation ability directly affects the application effect of translation results and is closely related to people's economic activities. However, there are grammatical errors in the English translation results, which make the computer translation results deviate and affect the output and judgment of the English translation results. Therefore, in previous studies, a large number of experts and scholars have proposed automatic identification methods for machine English translation errors, in an effort to reduce the impact of English translation errors on economic activities.

It has become the mainstream machine translation technology in the industry. Machine translation is basically available in general scenarios. Syntactic structure analysis refers to the process of judging whether the composition of a given input sentence conforms to the given grammar and analyzing the syntactic structure that conforms to the grammar rules. Generally speaking, syntactic structure analysis methods can be divided into two categories: rule-based analysis methods and statistics-based analysis methods. The rule-based syntactic structure analysis method uses hand-written grammatical rules to build a grammatical knowledge base, and at the same time uses conditional constraints and checking methods so that ambiguity and supergrammatical phenomena in sentences can be effectively dealt with. However, the workload is huge, and the knowledge bases in different fields are quite different, which is not conducive to the transfer of rules between fields. In addition, the analysis effect of this method is not good for long and difficult sentences. Therefore, it is necessary to study a more accurate automatic recognition method for machine English translation errors [21–23].

Information fusion is to integrate information from different sources and remove redundancy, and apply the obtained fusion information to various subsequent tasks. As shown in Figure 1, information fusion can be divided into three levels: data fusion, feature fusion, and decision fusion. With the rapid development of deep learning technology, the wide application of distributed processing technology, and the substantial improvement of computer processing capabilities, the advantages of feature fusion are becoming more and more obvious. Feature fusion is widely used in the field of images, and it is also applicable in the field of text similarity calculation. Fusion of features extracted in different ways is an important means to improve the performance of the model. The underlying features contain more detailed information and are more noisy. High-level features have stronger semantic information and less ability to perceive details. How to efficiently integrate the two, take advantage of it, and get rid of its dross, is the key to improving the performance of the model [24, 25].

Many researchers improve the performance of the entire model by fusing multilayer features. According to the order of fusion and model prediction, they are divided into early fusion and late fusion. Early fusion is to first fuse the features of multiple layers and then use the fused features to train the

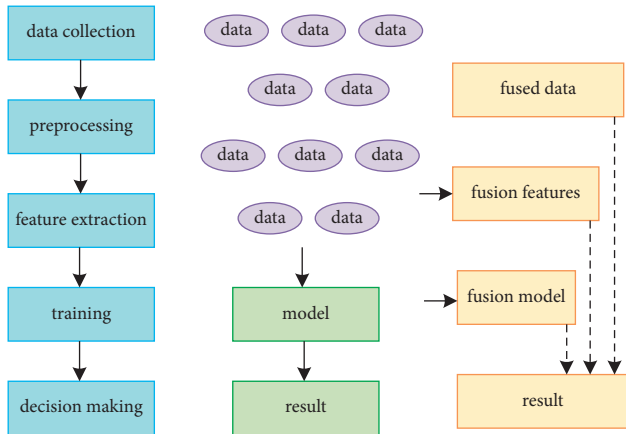


FIGURE 1: Multiple feature fusion levels.

model and only extract the features uniformly after complete fusion. This type of method is also called skip connection, that is, in a serial or parallel manner. The two classic feature fusion methods are as follows:

- (1) Concat fusion, that is, serial feature fusion, which directly connects the two feature vectors. If the dimensions of the two input feature vectors x and y are p and q , then the dimension of the output feature vector z is $p + q$;
- (2) The add mode fusion, that is, parallel feature fusion, combines two feature vectors into a complex vector. For the input features x and y , the fused feature is $z = x + iy$, where i is an imaginary unit. Late fusion is to improve model performance by combining the training results of different layers, start model training before completing the final fusion, and finally fuse multiple training results. There are two types of research ideas:
 - (1) The features are not fused, and the models are trained separately with multiscale features, and then the training results are synthesized
 - (2) The features are fused by pyramid network (PN), and the similarity is calculated after fusion

However, there are still many problems in data-driven neural machine translation technology, such as colloquial translation.

2. Multisignal Feature Fusion

In network flow classification, feature fusion can improve the accuracy. Low-rise and high-rise features each have their own strengths and weaknesses. Among them, the advantage of the former is that it contains rich information and has high discrimination ability; the disadvantage is that because it is very simple and does not undergo various transformations, it has lower semantics, and has a lot of noise at the same time; however, the latter is similar. On the contrary, in order to achieve the purpose of improving the classification performance, the most important point is to achieve the effective fusion of these two features.

Feature selection refers to picking out those parts from the existing features that are conducive to improving the prediction effect of the model, which is carried out before the training of the machine learning model. The problem of dimensional disaster may occur in stream classification tasks, and this link can effectively alleviate the dimensional disaster. Second, irrelevant features can be removed, reducing the difficulty of the classification task and making the model easier to understand. In addition, feature selection can remove irrelevant variables, minimizing the risk of overfitting. The feature selection process, subset search, evaluation methods, and common feature selection methods will be introduced below.

Feature fusion is very important to improve the accuracy of classification tasks. Feature fusion is divided into serial and parallel strategies. The former is to transform two or more features into one feature, which is suitable for fusing data of different dimensions. Since this fusion method does not need to unify the data dimension, it can effectively prevent the loss of information caused by the unification of the data dimension. The parallel strategy is to combine two feature vectors into a complex vector, suitable for models with the same output dimension. This method fuses different features without increasing the data dimension. Compared with the serial strategy, this method can effectively reduce the computational complexity. However, it needs to unify the data dimensions, which can lead to information loss. Errors, translation errors in professional fields, and low-resource language translation problems have not yet reached the ideal level under the current technical conditions, and the machine translation system will still output wrong translation results.

In image recognition, based on canonical correlation analysis, the transformation is achieved through the correlation between features, The resulting features will have higher correlations. However, the main limitation of CCA is that it ignores the relationships between classes. In order to overcome the shortcomings of CCA, a discriminant correlation analysis method is proposed, which not only retains the advantages of CCA but also increases the distinction between classes as much as possible. In order to reduce the loss of useful information to improve the classification accuracy, many methods have been proposed: fusion strategies through addition and L1 norm, fusion features from multiple feature extractors, through the combination of static and dynamic features, and through polynomial kernel functions feature fusion. A deep architecture-based feature generation model has been proposed, which improves the classification accuracy to a certain extent but requires a lot of data preprocessing time. While these feature fusion schemes provide better performance than single modality, they require separate feature extraction for each modality, resulting in longer processing time. Meanwhile, the fused features usually have high dimensionality, which increases the time overhead. In order to reduce the feature dimension, some people also convert the original features into semifeatures to maximize the mutual information between the transformed features and the target class. However, this approach may result in insufficient feature information generated. Feature

fusion synthesizes relevant information extracted from network traffic data and can be used to improve classification accuracy in many cases. Although some achievements have been made in applying feature fusion to network flow classification, there are still some problems, such as how to generate more discriminative features to improve classification accuracy and the increase of time complexity due to the introduction of feature fusion.

To sum up, how to generate or fuse more discriminative features to improve the performance of network traffic classification with lower time overhead is still a hot research topic.

2.1. Feature Extraction of Machine English Translation Information. In natural language processing, feature vectors are derived from text data and can reflect various linguistic characteristics of text data. This mapping from text data to concrete vectors is called feature representation and is done through a feature extraction model. Picking the right features is part of what makes a machine learning task successful, and deep neural networks alleviate the reliance on feature engineering, especially for linguistic data that exists in the form of a series of symbols, which need to be converted into a numeric vector using specific methods.

Features in text can be represented in a scalar or countable form. A scalar feature usually takes the value 0 or 1, depending on whether a certain condition occurs, such as when the word “cat” occurs at least 1 time in the text, the feature takes 1, otherwise, it takes 0. The value of a countable feature depends on the frequency of a given event, such as the number of times “cat” appears in the text as the feature value.

For a sentence, a paragraph or a piece of text, the features that can be considered are the number and word order of characters and words in the text. The bag-of-words (BOWs) process is a common process for extracting features from sentences and texts. This method abstracts words into basic elements by considering the number of each word as a feature. Since the existing translation error detection solutions have problems such as low efficiency, high cost, and waste of manpower, and have gradually been unable to meet the rapidly growing translation needs.

In addition, the statistical results of external information can also be combined to focus on those words that appear frequently in a given text but rarely appear outside. When using the BOW method, TF-IDF weights are mostly combined. For a piece of text d , which is part of the corpus, we denote each word w in d as a normalized result:

$$\frac{N_d(w)}{\sum_{w' \in d} N_d(w')} \quad (1)$$

The TF-IDF weights are

$$\frac{N_d(w)}{\sum_{w' \in d} N_d(w')} \times \log \frac{|D|}{|\{d \in D: w \in d\}|} \quad (2)$$

When considering words in sentences and texts, a directly observable feature of a word is its position in the

sentence, and words and characters surrounding it can also serve as features. The closer it is to the target word, the richer the amount of information the word has relative to distant words. For example, the N-gram model uses the first $n-1$ word sequences to predict the probability of word n , which includes local contextual features and text structural features, and is better than the bag-of-words model in practical applications.

In this research, the feature extraction algorithm of English translation will be integrated to make full use of parallel corpus. The extracted features are combined with the translation results to obtain the information features of the machine English translation. Through literature analysis, it can be found that machine translation can be divided into two parts, namely, the translation of the source language into the target language and the translation of the target language into the source language. The two translation processes are identical and share the word vector parameter. We set the source language statement to

$$A = \{a_1, a_2, \dots, a_n\}, \quad (3)$$

where a_i represents the word of the source sentence; the target sentence is

$$B = \{b_1, b_2, \dots, b_n\}, \quad (4)$$

where b_j represents the word embedding code of the target sentence; C represents the length of the source sentence; D represents the length of the target sentence. It is assumed that the encoder and decoder used in this translation are constructed as a neural network structure. The main function of the encoder is to encode the source sentence A into a fixed vector E , and at the same time decode E to obtain the target sentence D . The integrated translation process can be expressed as $P(B|A:\alpha)$, and the calculation process of the above conditional probability is obtained by using the multiplication rule, as shown in formula (5).

$$P(B|A:\alpha) = \prod_{i=1}^T p(b_j|a, b_1, b_2, \dots, b_{n-1}\alpha). \quad (5)$$

2.2. Multifeature Fusion Prediction of Machine English Translation. According to the extracted machine-English translation features combined with the translation automatic evaluation method, the machine-English translation results are predicted. Using the Pearson coefficient as a guiding factor, the translation results are initially analyzed, and the specific calculation process is set as formula (6):

$$l = \frac{oxy - oxoy}{\sqrt{dxdy}}. \quad (6)$$

Among them, o represents the mathematical expectation of the translation result; d represents the variance. In general, the value of this formula is -1 or 1 . When the calculation result has a high correlation, the value is close to 1 , otherwise, it is close to -1 . According to the above formula, taking into account the characteristics of machine translation, a penalty function is introduced in the

process of information prediction so as to ensure that the translation preference will not affect the translation result. Then, there is formula (7):

$$\mathfrak{F} - N = U^* \exp \sum_{i=1}^n \varepsilon_i \lg \text{precision}. \quad (7)$$

Among them, N is the number of penalty factors; ε_i is the flip vector; $st - 1$ is the influence vector of the hidden state at time $t - 1$ on the sentence; xhr is the word count vector of the decoder; kt is the encoder vector, and xhk is the encoder's word count vector. nt represents the step vector, and xhn represents the maximum encoding length vector, which is mainly used to limit the encoding steps; st represents the influence vector of the hidden state at time t on the sentence, kt represents the error vector, nt represents the error vector magnitude, and st is the machine English translation main cause of error.

We use this formula to predict the machine translation result and determine the correct rate of this result. At the same time, the information with a low accuracy rate is obtained as a training group for translation error recognition, and a corresponding support vector machine is constructed to make a secondary judgment on this part of the information. For the binary classification problem, in order to obtain the final reliable prediction result, the training set is set as

$$(z_i, y_i), i = 1, 2, \dots, n, z_i \in R^n, y_i \in \{\pm 1\}. \quad (8)$$

The classification plane can be expressed as

$$(q * z) + k = 0. \quad (9)$$

Here, k represents the penalized plane slope; q and z represent the length and width of the penalized plane, respectively.

According to formula (9), the samples are correctly distinguished, and the classification interval is maximized. The optimal classification result needs to meet the following conditions:

$$y_i [(q * z) + k] \geq 1. \quad (10)$$

Building a support vector machine based on formula (10), this problem can be optimized as

$$\left\{ \min 1/2\beta * \beta + G \sum_i \mathfrak{F}, \text{ s.t. } y_i [\varphi(q * z) + k] \geq 1 - \mathfrak{F}. \quad (11) \right.$$

Among them, G represents the cost coefficient in the classification process; $\varphi(\cdot)$ represents the nonlinear transformation function in the judgment process; i represents the slack variable function. According to this formula, the final judgment formula can be obtained as follows:

$$f(z) = \sum_{i=1}^m \eta_i y_i H(z_i, z) + k. \quad (12)$$

2.3. Design of Machine English Translation Error Recognition Algorithm. In order to effectively fuse the semantic features and context structure interaction features extracted by the

text semantic feature extraction module, text structure feature extraction module, and LSF feature extraction module, this paper studies different feature fusion methods, and compares different information fusion methods to help. According to the above setting results, a machine English translation error identification algorithm is designed to realize automatic identification of wrong translation. In order to make this algorithm feasible, the directed graph of wrong translation results is used as the main reference of the algorithm, and the directed graph of wrong translation is drawn, as shown in Figure 2.

In recent years, some researchers in this field have begun to try to utilize external knowledge sources as features, such as deep syntactic or deep semantic features. Word posterior probability features are based on system features and cannot provide sufficient knowledge of grammar or syntax. Therefore, it is considered to extract linguistic features such as syntax and semantics to improve the detection accuracy. This paper mainly uses two kinds of commonly used linguistic features: lexical features and syntactic features. A detailed description of these two types of features is given below.

For most of the words in the translation hypothesis, it is generally considered that the probability of a word sequence with high frequency and a sequence of part-of-speech tagging is higher than that of word sequence and part-of-speech tagging sequence with less frequency. Here, we consider the context information of the word and part-of-speech tagging, for each word and part-of-speech tagging, the first 2 and the last 2 and their own constitute feature vectors. Select a suitable feature fusion method to improve the accuracy of model calculation.

Syntactic knowledge is generally obtained after syntactic analysis of the source language or target language by a syntactic analyzer. When the analyzer cannot analyze the entire sentence, the problem word is ignored to find the connection of the remaining words to complete the syntactic analysis. Words that are ignored become words that are not connected to other words in the sentence, called null-link words. It is generally believed that these words that are not related to other words are more likely to have grammatical errors, and this grammatical information is used to define binary syntactic features. The evaluation is shown in Figure 3.

3. Experimental Design

In order to confirm that the automatic recognition method of machine English translation errors based on multifeature fusion proposed in this study has application value, an experiment was constructed to analyze the application effect of this method. The statistical machine translation system in the experiment of this paper is a phrase-based machine translation system: Moses system. This system is utilized to output 10,000 best translation hypotheses for each source language sentence during the decoding process of the test set, i.e., $N = 10000$. The direction of the translation language pair is Chinese to English translation. The training corpus of the translation model is the dataset provided by LDC, with a

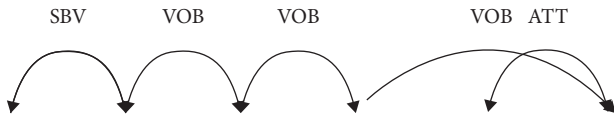


FIGURE 2: Directed graph of incorrect translation results.

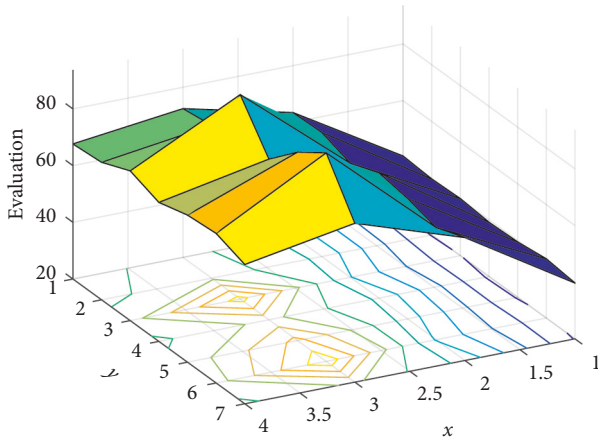


FIGURE 3: The evaluation.

total of 3,397,538 sentence pairs (mainly including Hong Kong news, FBIS, ISI Chinese-English network data, and Xinhua News). The language model adopts a five-element model, and its training corpus includes the English part of the above bilingual corpus and the Xinhua part of the English Gigaword, with a total of about 10 million sentences.

The development set of the machine translation system is the current set of the NIST machine translation evaluation task in 2006, with a total of 1,664 sentences, and each source language input corresponds to 4 reference answers. The test set is 1,082 sentences of NIST 2,005 current dataset and 1,357 sentences of NIST 2008 current data, and each sentence corresponds to 4 reference answers.

Data annotation: This paper leverages the WER criterion from the TER toolkit to determine the true classification results for each word in the translation hypothesis. First, the reference translation is found with the smallest edit distance from the translation hypothesis among the four reference translations as the benchmark, and the WER criterion is used to align the hypothetical translation with the reference translation. If the word in the translation hypothesis is consistent with the word in the same position in the reference translation marked as c , otherwise marked as i . The radar figure is shown in Figure 4.

In the NIST MT 2008 translation results generated by the SMT system, the 1-best translation is assumed to contain 38,587 words, of which 14,658 are labeled c and 23,929 are labeled i , whose samples are the “correct” class ratio (RCW) is 37.99%. The 1-best translation hypothesis of the NIST MT 2005 translation results contains 36,497 words, of which 15,179 are labeled c and 21,318 are labeled i , and the proportion of classes whose samples are “correct” (RCW) is 41.59%. Dev and Test Sets: The NIST MT 2008 dataset (1 357 sentences) is used as the development set for maximum entropy model parameter training in the

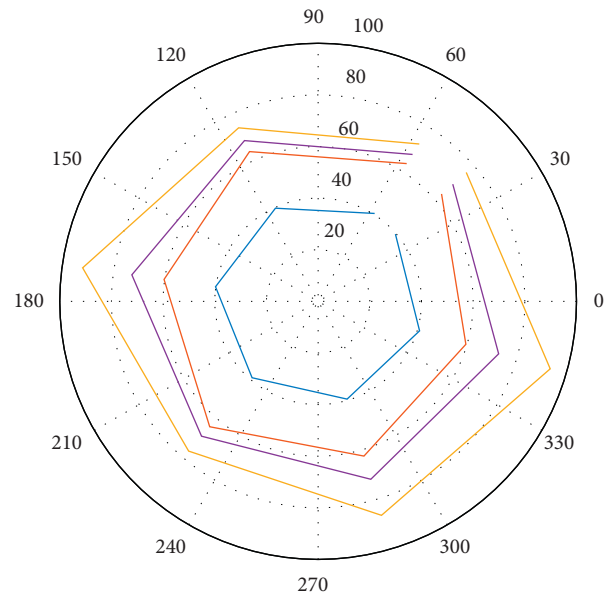


FIGURE 4: Radar figure.

translation error detection task, and the NIST MT 2005 (1,082 sentences) dataset is used as the test set for the classification task.

The maximum entropy classifier is used to conduct independent feature classification experiments on three typical WPP features, three linguistic features, and extracted source word features, and analyze and compare the results. The results of the classification experiments are shown in Figure 5.

In Figure 5, WPP_Dir represents WPP based on fixed position; WPP_Lev represents WPP based on Levenshtein alignment, and WPP_Win represents WPP based on target position window. The word feature is represented by Wd, and the source word feature is represented by Source.

The CER of the baseline system is 41.59%. The following conclusions can be drawn from the data in Figure 6.

- (1) When the three typical WPP features are used independently, the CER value can be reduced to a certain extent, and the calculation method based on the target position window obtains a lower CER result than the other two methods because of its flexibility. WPP based on Levenshtein alignment performs best in terms of F-values.
- (2) Linguistic features perform better than standalone WPP features (except for Link). The word features can help the system find the error more accurately. The grammatical feature link has the highest recall rate, indicating that this feature can help the system find more errors. The data in Figure 6 also reveal the contribution of linguistic features to error detection, indicating that linguistic features can effectively reduce the classification error rate and improve the error prediction ability.
- (3) As far as CER is concerned, the source word features are reduced 4.59%. It proves the effectiveness of the source-side word features. It can be

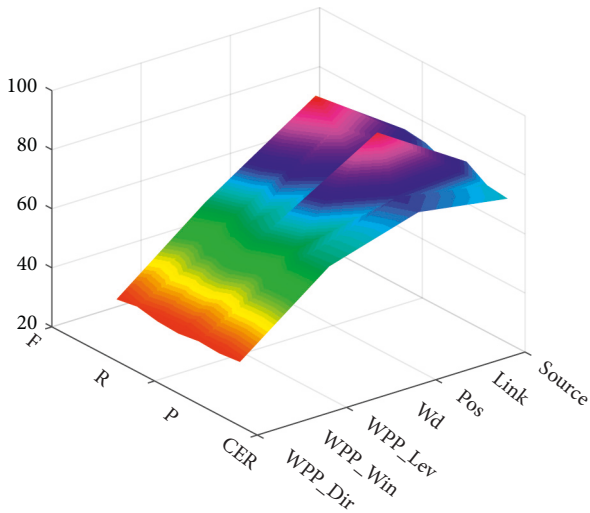


FIGURE 5: Results of individual feature for translation error detection.

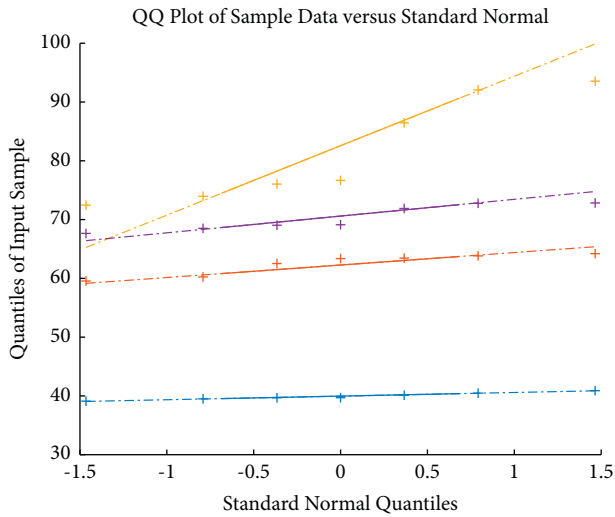


FIGURE 6: Prediction variation.

considered to extract richer source-side information as features and add them to the classifier to improve the performance of error detection and classification. The prediction variation is shown in Figure 6, which means the validation of the proposed when processing the machine English translation errors.

In the process of this experiment, the experimental platform is set as Windows and Linux system. In this system, the collection and processing of original translation information and extended information are completed, and the experimental part will be completed in Linux system. During the experiment, JAVA is used as the experimental control language, and the processing of files and the output of experimental results are controlled by this language. At the same time, we set the rules for merging experimental results, process the experimental results, and output the results.

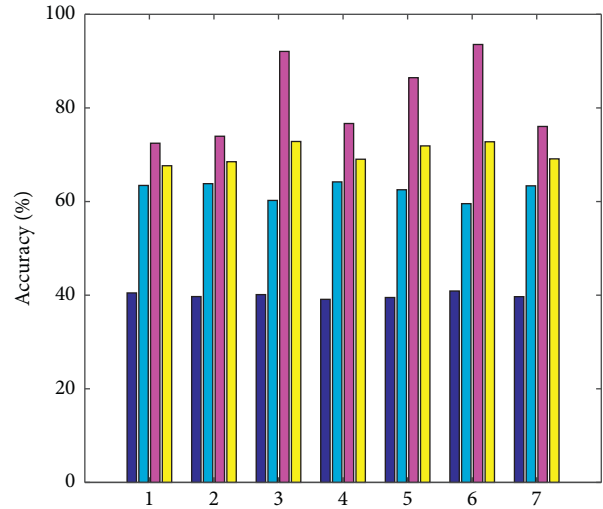


FIGURE 7: Accuracy.

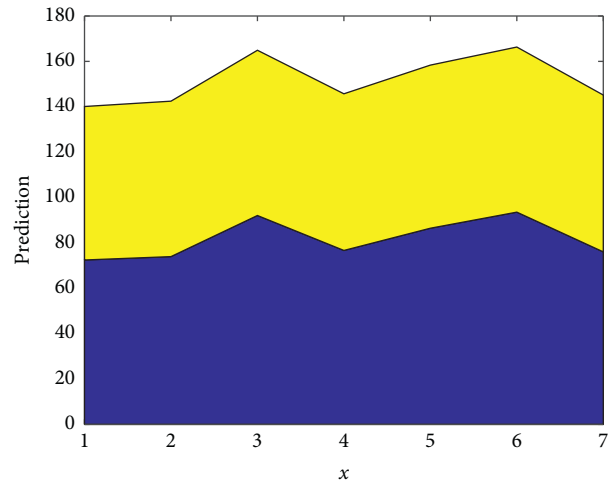


FIGURE 8: Prediction.

The training data in the experiment mainly come from a laboratory database. The training dataset contains 5,000 wrong sentences and 5,000 corresponding correct sentences. These sentences are checked by all English-speaking staff who manually mark grammatical errors and correct each sentence. A mistake was obtained after combining this part of the information, and 10 experimental datasets were constructed, as shown in Figure 6.

The collected translation information is divided according to the content in Figure 6, and the word vector is trained at the same time. We use the Word2vec tool to train the translation information, set the vocabulary vector dimension of the translation information to 1024, set the window size to 10, use the negative sampling optimization algorithm to set the number of translation information samples to 10, and set the number of iterations to 20 times. In the experimental preparation stage, in order to ensure the reliability of the experimental results and reduce the error of the experimental results, the training set was trained using the translation information template preset in the previous

research, and the training set was marked so as to divide and process the experimental data. The accuracy is also compared in Figure 7. Besides, the prediction is shown in Figure 8.

4. Conclusion

Aiming at the current English translation results, this paper proposes a new method for automatic recognition of translation errors, which has been proved to have certain practical effects through experiments. This time, the focus of the research is based on the accuracy of recognition, and there is no optimization for other fields. For this reason, other parts need to be analyzed in the follow-up research, and the shortcomings of this method should be improved and optimized to improve the translation effect and provide help for the development of machine translation technology.

Although the text similarity calculation model that combines various features proposed in this paper has achieved good performance, the model efficiency needs to be further optimized. In the future, the distributed computing framework can be combined with multiple graphics cards for parallel processing to improve the entire model of operating efficiency.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] H. Li, "Virtual interaction algorithm of cultural heritage based on multi feature fusion," *Journal of Computational Methods in Science and Engineering*, vol. 22, no. 1, pp. 333–347, 2022.
- [2] K. Miao, J. Ma, Z. Li, Y. Zhao, and W. Zhu, "Research on multi feature fusion perception technology of mine fire based on inspection robot[J]," *Journal of Physics: Conference Series*, vol. 1955, no. 1, pp. 76–86, 2021.
- [3] Y. Chen, Z. Wu, B. Zhao, C. Fan, and S Shi, "Weed and corn seedling detection in field based on multi feature fusion and support vector machine," *Sensors*, vol. 21, no. 1, p. 212, 2020.
- [4] R. Priyadharshini and R. Ramkumar, "Faster R-CNN network based on multi feature fusion for efficient face detection[J]," *International Journal of Recent Technology and Engineering*, vol. 9, no. 2, pp. 495–499, 2020.
- [5] F. Li and C. Li, "Rural financial development level, rural finance supply and rural economy-an empirical study based on panel threshold model," in *Proceedings of 2018 8th ESE International Conference on Management Science, Education Science and Human Development*, pp. 29–33, Springer Nature, Berlin, Germany, June 2018.
- [6] C. Li, G. Jiang, and J. M. Dewaele, "Understanding Chinese high school students' foreign language enjoyment: validation of the Chinese version of the foreign language enjoyment scale," *System*, no. 76, pp. 183–196, 2018.
- [7] Y. Zhu and Y. Jiang, "Optimization of face recognition algorithm based on deep learning multi feature fusion driven by big data," *Image and Vision Computing*, vol. 104, no. 1, p. 104023, 2020.
- [8] X. Huang, "A pedestrian tracking method based on adaboost and feature fusion," *IOP Conference Series: Earth and Environmental Science*, vol. 252, no. 3, Article ID 32137, 2019.
- [9] R. Xu, J. Han, W. Qi, J. Meng, and H. Zhang, "Railway fastener image recognition method based on multi feature fusion[J]," *IOP Conference Series: Materials Science and Engineering*, vol. 397, no. 1, pp. 43–54, 2018.
- [10] Y. Feng and H. Zhou, "Research on medical image classification and recognition based on multi feature fusion [J]," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9, no. 12, pp. 113–122, 2016.
- [11] Z. Deng, S. Song, and H. Tan, "A new analytical framework for urban size and policies," *Applied Economics Letters*, vol. 28, no. 18, pp. 1610–1619, 2021.
- [12] H. Su, G. Han, L. Li, and H. Qin, "The impact of macro-scale urban form on land surface temperature: an empirical study based on climate zone, urban size and industrial structure in China[J]," *Sustainable Cities and Society*, vol. 74, pp. 33–45, 2021.
- [13] Y. Zhan, C. Mengyao, C. Yuefen, L. Zhaoqian, L. Yao, and Li Kedi, "An automatic recognition method of fruits and vegetables based on depthwise separable convolution neural network[J]," *Journal of Physics: Conference Series*, vol. 1871, no. 1, pp. 1–9, 2021.
- [14] W. Wu, C. Li, Z. Wu, and Y Yin, "Automatic recognition method of primary path for an anastomosing river based on its typical features," *Transactions in GIS*, vol. 25, no. 3, pp. 1577–1598, 2021.
- [15] W. Huang, N. Li, Z. Qiu, N. Jiang, B. Wu, and B Liu, "An automatic recognition method for students' classroom behaviors based on image processing," *Traitement du Signal*, vol. 37, no. 3, pp. 503–509, 2020.
- [16] G. Zhang, D. Cheng, Y. Hou, Z. Li, and L Zhong, "Study on automatic recognition method of Continental Shale Sandy laminae based on electrical imaging image," *Journal of Physics: Conference Series*, vol. 1549, no. 2, 2020.
- [17] K. Hong Joo and Ko Ji Yeong, "A comparative study on the housing conditions by urban size in local small-medium cities," *Journal of the Korean Housing Association*, vol. 30, no. 1, pp. 77–85, 2019.
- [18] P. Wang, K. Gu, J. Hou, and B. Dou, "An automatic recognition method for airflow field structures of convective systems based on single Doppler radar data," *Atmosphere*, vol. 11, no. 2, pp. 142–232, 2020.
- [19] J. Zhang, R. Jiang, B. Li, and N Xu, "An automatic recognition method of microseismic signals based on EEMD-SVD and ELM," *Computers & Geosciences*, vol. 133, no. C, p. 104318, 2019.
- [20] D. Castells Quintana, V. Royuela, and P. Veneri, "Inequality and city size: an analysis for OECD functional urban areas," *Papers in Regional Science*, vol. 99, no. 4, pp. 1045–1064, 2020.
- [21] N. Pourdamghani and K. Knight, "Neighbors helping the poor: improving low-resource machine translation using related languages," *Machine Translation*, vol. 33, no. 3, pp. 239–258, 2019.
- [22] F. Fang, "The construction of college English translation teaching mode in the information technology environment

- [J],” *International Journal of Social Sciences in Universities*, vol. 3, no. 3, pp. 1–8, 2020.
- [23] Y. Gao, “Normative research on English translation of tourist attractions in inner Mongolia[J],” *The Frontiers of Society, Science and Technology*, vol. 2, no. 11, pp. 1866–1886, 2020.
- [24] C. Mai, “A study on the current situation and countermeasures of college English translation teaching[J],” *International Journal of New Developments in Education*, vol. 2, no. 4, pp. 1–7, 2020.
- [25] Pu Zhou and Z. Jiang, “Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin[J],” *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512–524, 2016.

Research Article

Optimization of Reservoir Flood Control Operation Based on Multialgorithm Deep Learning

Bowen Xue ^{1,2}, **Yan Xie**¹, **Yanhui Liu** ^{2,3}, **Along Li**⁴, **Daguang Zhao**^{1,5} and **Haipeng Li**^{6,7}

¹School of Water Conservancy, North China University of Water Resources and Electric Power, Zhengzhou, Henan 450046, China

²Yellow River Institute of Hydraulic Research, Zhengzhou, Henan 450003, China

³College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing, Jiangsu 210098, China

⁴Yellow River Engineering Consulting Co., LTD, Zhengzhou, Henan 450003, China

⁵Jilin Province Water Conservancy and Hydropower Survey and Design Institute, Changchun, Jilin 130021, China

⁶College of Agricultural Science and Engineering, Hohai University, Nanjing, Jiangsu 210098, China

⁷Technical Advisory of PRWRC (Guangzhou) Co., Ltd, Guangzhou 510000, China

Correspondence should be addressed to Yanhui Liu; liuyanhui@hhu.edu.cn

Received 2 April 2022; Revised 20 April 2022; Accepted 26 April 2022; Published 10 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Bowen Xue 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 development of China's social economy, it is the most important task for the water conservancy industry to make use of the existing water conservancy engineering measures to carry out the research on river basin flood control dispatching. Large-scale joint operation of river basins usually needs to consider meteorological and hydrological conditions, historical flood data, multireservoir engineering conditions, and multiple flood control targets, which is a complex decision-making problem. Therefore, electing the optimal operation model of reservoir flood control optimization is very important. In this paper, Luanhe River Basin is taken as the research area, and three kinds of constraints, namely, water balance constraint, reservoir flood control capacity constraint, and water release decision constraint, are set to construct the flood control optimization model. Taking the minimum square of the sum of reservoir discharge and interval flood discharge as the objective function, genetic algorithm (GA), particle swarm optimization (PSO), Spider swarm optimization (SSO), and grey wolf optimization (GWO) are introduced into flood control optimal operation to seek the minimum value of objective function, and the results are compared and analyzed. Through the analysis of optimization results, the optimization ability and convergence effect of grey wolf optimization algorithm are better than those of genetic algorithm and particle algorithm, and the results are more stable than those of spider swarm algorithm. It has a good model structure and can make full use of the results of three wolf groups for optimization. Through the analysis of scheduling results, the results of genetic algorithm and particle swarm optimization algorithm are similar, while those of spider swarm optimization algorithm and grey wolf optimization algorithm are similar and slightly better than those of the first two. Moreover, the search range of grey wolf optimization algorithm for solving long sequence problems is wider and the calculation time is shorter. Therefore, the grey wolf optimization algorithm can be applied to solve the flood control operation optimization model of Panjiakou Reservoir Group.

1. Introduction

With the rapid development of China's social economy, the impact of human activities on the environment is aggravated, and the economic losses caused by floods account for the first of all kinds of natural disasters in China. How to make use of the existing water conservancy project measures and scientifically carry out the research on flood control

dispatching in river basins is the most important task of disaster prevention [1] and mitigation in water conservancy industry at present. When guiding the flood control operation of reservoir groups, it still depends more on the experience of decision-making level. There are still some problems when the existing mathematical models and scientific theories are applied to the actual operation [2, 3]. With the development of modernization, more and more

reservoirs with flood control functions have been built in large and medium-sized river basins, and the research on joint optimal operation of reservoir groups has become one of the hot topics. In the past reservoir operation, some scholars [4] showed that the joint operation model of cascade reservoirs is constructed, considering the uncertain inflow from the upstream of reservoirs, and the Copula function is used for modeling, which has achieved remarkable results. On this basis, particle swarm optimization is adopted [5]. Increase the calculation speed and improve the algorithm [6]. Improving the global optimization ability makes it possible to use mathematical analysis method to formulate a scientific dispatching scheme for reservoir dispatching and give better play to the role of water conservancy projects.

The intelligent optimization algorithm is mostly used in the optimization of reservoir flood control optimal operation model. With the deepening of research, the complexity of reservoir flood control optimal operation is gradually excavated, and the mathematical model based on this is nonlinear, multiobjective, and high-dimensional [7]. In order to improve the efficiency and accuracy of model solution [8], traditional optimization methods often fail to meet the complex requirements of new problems [9]. Therefore, more and more optimization algorithms have been developed and utilized by scholars in flood control optimal operation, such as genetic algorithm [10, 11]. The algorithm based on evolutionary thinking uses natural evolutionary rules to achieve iterative optimization. Based on spider swarm algorithm [12], by simulating the social behavior of social spiders in the natural environment, the optimization target is drawn up as the real object on the spider web, and the female and male spiders are drawn up as two optimization methods and search for "food" on the spider web according to their different standards. Wolves algorithm [13] mainly simulates the behavior of wolves searching for prey, surrounding prey, and attacking prey [14]. On the basis of conventional dispatching theory, the combined optimal dispatching of reservoirs uses advanced intelligent optimization algorithm and system science concept to solve the optimal strategy which meets the operation requirements of reservoirs in the basin [15], so as to obtain the benefits that are difficult to realize when a single reservoir is operated. Large-scale joint operation of river basins usually needs to consider meteorological and hydrological conditions, historical flood data, multireservoir engineering conditions, and multiple flood control targets, which is a complex decision-making problem [16]. According to the characteristics of Luanhe River basin and flood control requirements, this paper discusses the optimal dispatching model of reservoir flood control optimization based on multialgorithm deep learning.

The latitude and longitude range of Luanhe River Basin is $115^{\circ}30' \sim 119^{\circ}45'$ east longitude, $39^{\circ}10' \sim 42^{\circ}40'$ north latitude, the northern end to the southern edge of Inner Mongolia Plateau, the intersection with Bohai Sea in the south, Chaobai and Jiyun Canal in the west, Liaohhe River in the east, and the basin area of 44750 km^2 . Panjiakou, Daheiting, and Taolinkou reservoirs are the three most

important large-scale water conservancy projects in Luanhe River Basin. Among them, Panjiakou and Daheiting reservoirs are the main flood control projects on the main stream of Luanhe River, and the controlled basin area accounts for more than 75% of the total area of Luanhe River. Taolinkou Reservoir is the main flood control project on Qinglong River, a tributary, and these three reservoirs control 90% of the area of Luanhe River. The birthplace of Luanhe River is mainly grassland, with a gentle and open terrain, with an average elevation of 1350 m. The slope of riverbed is about 0.0005. Below Panjiakou Reservoir, the river is 200~500 m wide. After flowing through Sangyuan Canyon, the river widens greatly. After going down to Jingshan Railway Bridge in Luanxian County, the river flows into the plain [17–19]. Luanhe River Basin belongs to a relatively independent water system, and its source flows into the lower main stream and then into the ocean. Considering the length of the river in the lower reaches of the plain terrain area, there is no interval for flood storage, so flood prevention mainly depends on precaution and small dam. See Figure 1 for the schematic diagram of Luanhe River Basin.

2. Materials and Methods

2.1. Model Building

2.1.1. Objective Function. Setting an appropriate objective function according to the current situation of the basin and the characteristic value of the reservoir is one of the keys to determine the results of flood control optimal operation. Aiming at Panjiakou Reservoir Group in Luanhe River Basin, the joint optimal operation model takes Luanxian Station in the lower reaches of Luanhe River as the control point, and the flow of the control point is mainly composed of three parts: the discharge of Panjiakou Reservoir after flood calculation, Taolinkou Reservoir after flood calculation, and flood in Pan-Tao-Luanhe section, and the discharge of Luanxian Station is a combined discharge process of the above three parts.

There are many objectives for optimal operation of reservoir group joint flood control. At present, the maximum peak clipping criterion and the minimum disaster duration criterion are widely used. In this paper, the maximum peak clipping criterion is used as the objective function of reservoir group joint flood control operation, and the peak clipping rate is used as the evaluation criterion of operation results. The objective function is as follows:

$$F = \min \sum_{t=1}^n (q_{\text{pan } t+\Delta t_1} + q_{\text{tao } t+\Delta t_2} + Q_{\text{qu } t+\Delta t_3})^2. \quad (1)$$

In the formula, t is the serial number of time slots, and N is the total number of scheduled time slots; Δt_1 is the propagation time (h) from Panjiakou Reservoir discharge to the control section; Δt_2 is the propagation time (h) from the discharge of Taolinkou Reservoir to the control section; Δt_3 is the propagation time (h) from flood in Pan-Tao-Luanhe section to control section; $q_{\text{pan } t}$ is water discharge after

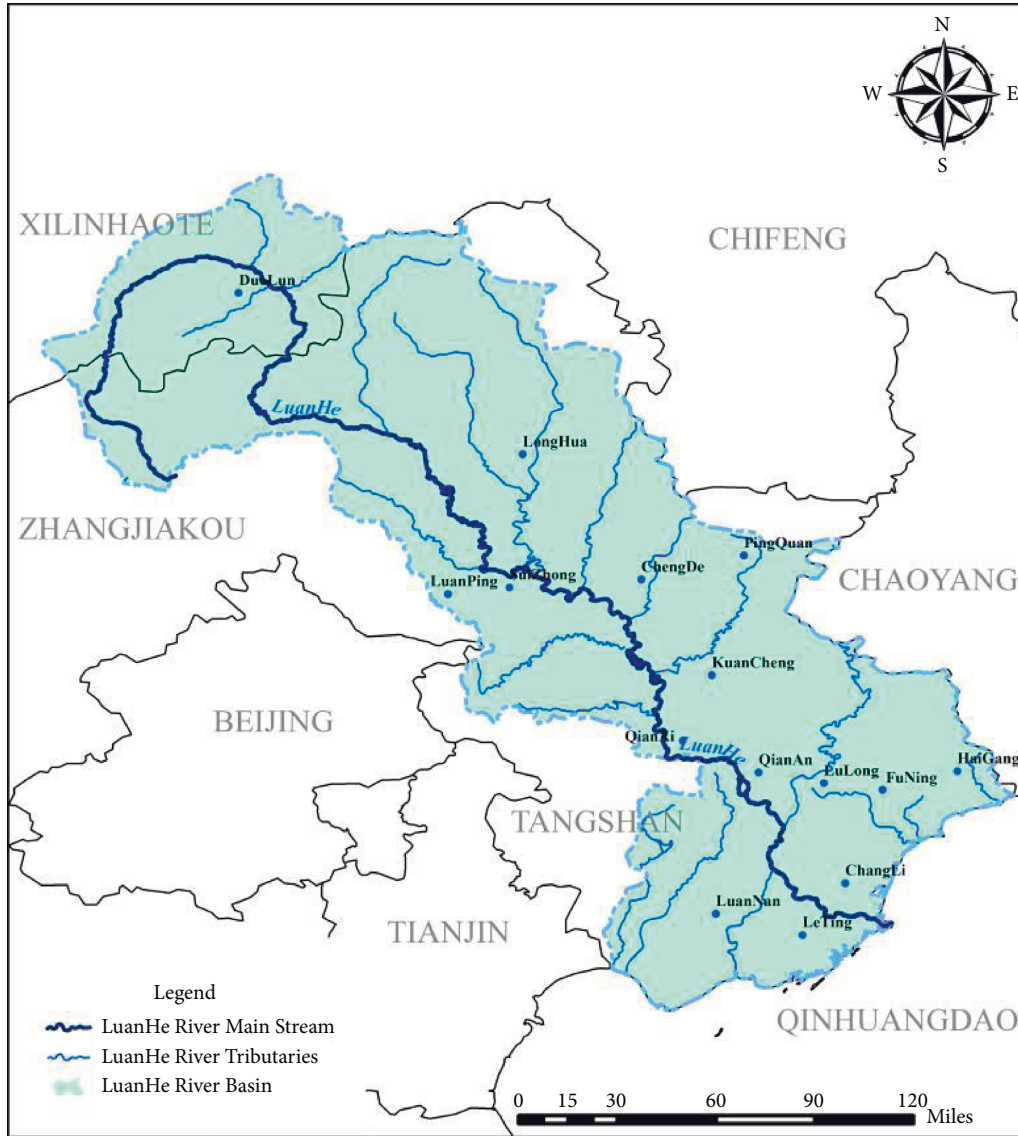


FIGURE 1: The drainage map of Luanhe River Basin.

regulation and storage by Panjiakou Reservoir (m^3/s); q_{taof} is water discharge after regulation and storage by Panjiakou Reservoir (m^3/s); Q_{qut} is Pan-Tao-Luanhe section flood discharge (m^3/s).

2.1.2. Model Constraints. After the Luanhe River flood control system was completed, it experienced several floods in 1962, 1989, 1994, 2012, etc. Considering the problems existing in flood control in Luanhe River Basin, three kinds of constraint targets for reservoir flood control were set. They are water balance constraint, reservoir flood control capacity constraint, and water release decision constraint.

Water balance constraint:

$$v_t = v_{t-1} + (q_{i1} - q_{i2})\Delta t. \quad (2)$$

Flood control capacity constraints:

$$\sum_{i=1}^{n'} (q_{i1} - q_{i2})\Delta t \leq v_3. \quad (3)$$

Decision-making constraints:

$$q_{i1} \leq q_{i2}, q_{i2} \leq q(z_t, b_t). \quad (4)$$

Type: v_t, v_{t-1} is reservoir capacity at the beginning and end of the period (m^3); Δt is length of time period (h); z_t is the water level of the reservoir at time (m); b_t is flood discharge capacity of the reservoir (m^3/s).

2.2. Comparison and Selection of Model Solution Methods. There are two key problems in the study of optimal operation of reservoir group joint flood control: 1. How to build a mathematical model that accords with the reality of the research object; 2. Find an effective algorithm to solve the model.

Genetic algorithm (GA) [20], particle swarm optimization (PSO) [21], spider swarm optimization (SSO) [21], and grey wolf optimization algorithm (GWO) [22], which are among the intelligent optimization algorithms, are

applied to solve the Panjiakou Reservoir Group's joint flood control optimal operation model, and the efficiency and optimal operation results of each algorithm are compared and analyzed.

2.2.1. Genetic Algorithm. Genetic algorithm takes the evolution of organisms as its basic idea, and each chromosome is the outflow sequence. It uses real numbers to encode and simulate the crossover, mutation, and selection of chromosomes, but it has the characteristics of premature convergence and difficulty in finding the optimal solution. The core of genetic algorithm is crossover, mutation, and selection. The crossover operator selects the dominant individual based on the fitness of the individual to combine the genes to form a new individual, which can better retain the dominant genes. The mutation operator can make the value of individual gene change randomly, so that the algorithm can generate a new optimal individual and get the global optimal solution. Selecting the fitness is based on the objective function. Choose by roulette, etc.

(1) *Crossover.* The crossover operation used in this paper is the single-point crossover method; that is, a crossover point is randomly selected, and the parts before and after the crossover point of two individuals are combined to form a new individual. G_1 and G_2 are two discharge sequences, respectively, which produce a crossing point P in G_1 and G_2 . The sequence before the crossing point P is unchanged, and the sequence after the crossing point is interchanged.

Before: $\mathbf{G}_1 = \{x_1, x_2, \dots, p, x_{n-m}, \dots, x_n\}$, $\mathbf{G}_2 = \{y_1, y_2, \dots, p, y_{n-m}, \dots, y_n\}$

After crossing: $\mathbf{G}_{11} = \{x_1, x_2, \dots, p, y_{n-m}, \dots, y_n\}$, $\mathbf{G}_{21} = \{y_1, y_2, \dots, p, x_{n-m}, \dots, x_n\}$

(2) *Variation.* Variation is based on the variation operator changing a certain value in the discharge sequence to obtain a new discharge sequence. G is the discharge sequence, $G = \{x_1, x_2, \dots, p, x_{n-m}, \dots, x_n\}$, which changes to q at the position of factor p , forming $G^2 = \{x_1, x_2, \dots, q, x_{n-m}, \dots, x_n\}$. If the fitness function of the formed sequence is better, the variation will be retained and passed on to the next generation.

(3) *Selection.* Selection is to pass on a better individual to the next generation. The roulette method is used in this paper to select the individual with the smallest fitness function. The fitness of an individual is converted into probability, which is the probability that the individual will be passed on to the next generation. The individual is selected based on the calculated probability and random number.

$$P_i = \frac{1/F(i)}{\sum_{i=1}^n 1/F(i)} \quad (i = 1, 2, \dots, n), \quad (5)$$

$$P_i = P_{i-1} + P_i,$$

where $F(i)$ is fitness of the discharge sequence, the number of samples of n -discharge sequence; P_i is the probability that

Inputs: objective function $f(x)$, search the upper and lower limits UB and LB of the space, and set the population individual number N and the maximum iteration number T ;

1: randomly initialize the population position $X_i (i=1, 2, \dots, N)$. Where $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$, d represents the dimension, and $t=0$;

2: While $t < T$ do

3: evaluate the fitness of the population and carry out roulette wheel selection operation;

4: cross-operate the two selected chromosomes, and replace the chromosome with the worst fitness in the population with the new chromosome (sub-region search);

5: randomly change a gene fragment of the selected chromosome (unknown region search);

6: generate a population with a new fitness value;

14: $t=t+1$;

15: endwhile

Outputs: the optimal individual in the final population and its fitness value.

FIGURE 2: The pseudocode of genetic algorithm.

the fitness of the i -th sequence accounts for all sequences; P_i is the cumulative sum of the probabilities of the i -1st sequence to the i -th sequence.

See the detailed development process of genetic algorithm and the pseudocode of genetic algorithm as shown in Figure 2.

2.2.2. Particle Swarm Optimization. Particle swarm optimization (PSO) is based on the simulation of the behavior of birds flying and preying on groups, and it searches for the best individual based on the evolutionary selection of the best individual of groups and individuals. In the search range, each particle has two characteristics: speed and position, and the particle constantly updates its position and speed by learning the optimal position of the particle itself and the optimal position of the group, so as to search for the best individual. The core of particle swarm optimization algorithm is the update and iteration of speed and position.

Update and iteration of speed:

$$V_i(j+1) = \zeta V_i(j) + \delta_1 \rho_1 (pB_i(j) - X_i(j)) + \delta_2 \rho_2 (gB(j) - X_i(j)). \quad (6)$$

Update and iteration of location:

$$X_i(j+1) = X_i(j) + V_i(j+1), \quad (7)$$

where V_i is the velocity vector of the i th particle, with m elements $(v_i^1, v_i^2, \dots, v_i^m)$; X_i is the position vector of the i th particle, with m elements $(x_i^1, x_i^2, \dots, x_i^m)$; ζ is inertia weight; j is number of iterations; δ_1, δ_2 are learning factors, used to characterize the importance of pB_i, gB ; ρ_1, ρ_2 are random numbers on the interval; pB_i, gB are the best position to separate particle i from group particles.

See detailed development process of particle swarm optimization and pseudocode of particle swarm optimization as shown in Figure 3.

2.2.3. Spider Swarm Algorithm. Spider swarm optimization algorithm is to simulate the information exchange, mating, and selection of male and female spiders. Spiders transmit information through vibration on the spider web. Spiders are divided into male and female. The male and female spiders transmit information based on the vibration model.

Inputs: objective function $f(x)$, search the upper and lower limits UB and LB of the space, and set the number n of individuals in the population group and the maximum iteration number T ;
1: randomly initialize the population position X_i ($i=1, 2, \dots, N$) and its velocity vector V_i . $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$, and D represents the dimension; Initialization factors δ_1, δ_2 and inertia weight $\zeta, t=0$;
2: While $t < T$ **do**
3: for each individual in the population
4: evaluate individuals and update the best information of individual history;
5: updating the global optimal particle information;
5: endfor;
6: update each particle position according to equations (5.9) and (5.10);
7: $t=t+1$;
8: endwhile
Outputs: the position X_{rabbit} and energy E_{rabbit} of the rabbit are the best points and optimal values.

FIGURE 3: The pseudocode of particle swarm optimization algorithm.

The dominant male spiders mate with the female spiders and protect the whole population for foraging. Information exchange and mating behavior of male and female spiders are the core of spider algorithm.

(1) *Algorithm initialization.* Initialization of spider group, including the number of male and female spiders, position, and weight initialization is as follows:

$$N_f = \text{floor}[(0.9 - \text{rand} \cdot 0.25) \cdot N], N_m = N - N_f,$$

$$s(i) = LB + \text{rand}(1, D) \cdot UB, \quad (8)$$

$$\omega(i) = \frac{J(X(i)) - \text{worst}(X)}{\text{best}(X) - \text{worst}(X)}.$$

In the above formula, n_f and n_m are the numbers of female and male spiders, n is the population review, rand is the random number generated by uniform distribution $[0,1]$,

$$f_i(j+1) = \begin{cases} f_i(j) + \delta_1 \times v_{i,c}(P_c - f_i(j)) + \delta_2 \times v_{i,b}(P_b - f_i(j)) + \delta_3 \times (\text{rand} - 0.5), & \delta_4 < PF \\ f_i(j) - \delta_1 \times v_{i,c}(P_c - f_i(j)) - \delta_2 \times v_{i,b}(P_b - f_i(j)) + \delta_3 \times (\text{rand} - 0.5), & \delta_4 \geq PF \end{cases} \quad (10)$$

$$PF = \exp(-pf_i).$$

Type: $\delta_1, \delta_2, \delta_3, \delta_4$ are random numbers in the interval of $\text{Rand}[0,1]$; j is the number of iteration cycles; P_c is individuals with the smallest distance and larger weight from spiders I ; P_b is individuals with the smallest distance and larger weight from spiders I ; P_f is probability factor used to control the attraction and repulsion behavior of spiders.

Inputs: objective function $J(x)$, search the upper and lower limits UB and LB of the space, and set the number n of spiders and the maximum iteration number T ;
1: Determine the number of female and male spiders according to formula (5.11), and initialize the population positions f_i ($i=1, 2, \dots, N_f$), m_i ($i=1, 2, \dots, N_m$) of different sexes. D is the dimension of the problem, the fitness and individual weight are updated, and the mating radius is initialized according to formula (5.18), with $t=0$;
2: While $t < T$ **do**
3: Update the position of female spider according to formula (5.15) and formula (5.16)
4: Update the position of the male spider according to formula (5.17)
5: Generate new individuals according to roulette method;
6: update that fitness value and individual weight of each individual;
7: $t=t+1$;
8: endwhile
Outputs: Select the best individual and fitness value, i.e., the best advantage and the best value, in female spider colony and male spider colony.

FIGURE 4: The pseudocode of particle social-spider optimization algorithm.

LB and UB are the upper and lower limits of the search space, and floor is the rounding function, ω is the individual weight, $J(s(I))$ is the fitness value, $\text{best}(s)$ is the optimal fitness value, and $\text{worst}(s)$ is the worst fitness value.

(2) *Information exchange.*

(a) The mathematical model of the vibration perception ability of individual J to individual I is as follows:

$$v_{i,j} = \omega_j \times e^{-\frac{r}{d_{i,j}}}. \quad (9)$$

Type: $v_{i,j}$ is information on the female nearest to spider $I(v_{i,f})$, the information of the best spider nearest to spider $I(v_{i,b})$, and information of individuals closest to spider I and with larger weight ($v_{i,c}$); ω_j is individual weight; $d_{i,j}$ is the distance between I and j of spiders.

(b) Vibration patterns of female spiders:

(c) Vibration patterns of male spiders: male spiders can automatically be identified and gathered, which can be divided into dominant male spiders and non-dominant male spiders. Dominant spiders can attract female spiders, and nondominant spiders move towards dominant spiders. The movement of non-dominant spiders to dominant spiders can be expressed as follows:

$$m_i(j+1) = \begin{cases} m_i(j) + \delta_1 \times v_{i,f}(P_f - m_i(j)) + \delta_3 \times (\text{rand} - 0.5), & \text{if } (\omega_{N_{f+i}} > \omega_{N_{f+m}}), \\ m_i(j) + \delta_1 \times \left(\frac{\sum_{h=1}^{N_m} m_h^j \times \omega_{N_{f+h}}}{\sum_{h=1}^{N_m} \omega_{N_{f+h}}} \right), & \text{if } (\omega_{N_{f+i}} \leq \omega_{N_{f+m}}). \end{cases} \quad (11)$$

In the above formula, P_f is the female nearest to the dominant male; $\omega_{N_{f+m}}$ is the weight of spiders in the middle; $\sum_{h=1}^{N_m} m_h^j \times \omega_{N_{f+h}} / \sum_{h=1}^{N_m} \omega_{N_{f+h}}$ is middle position of male spider; $\omega_{N_{f+i}} > \omega_{N_{f+m}}$ is the distinguishing conditions of dominant male spiders; $\omega_{N_{f+i}} \leq \omega_{N_{f+m}}$ is the discriminant conditions of nondominant male spiders.

(3) *Mating Behavior*. Male dominant spiders will mate with female spiders within the mating radius. If there are more female spiders within the mating radius, roulette will be used to produce new spiders.

Mating radius can be calculated by the following formula:

$$r = \frac{\sum_{k=1}^n (P_k^h - P_k^l)}{2n}. \quad (12)$$

Type: P_k^h, P_k^l are the upper and lower limits of the k -th dimensional variable of the discharge sequence.

See the detailed development process of spider swarm algorithm and the pseudocode of spider swarm algorithm as shown in Figure 4.

2.2.4. Grey Wolf Optimization Algorithm. The algorithm is based on the division of labor and information interaction in wolves' predation to realize the search of the optimal solution. Wolves are divided into α -wolf, β -wolf, δ -wolf, and other wolves. α -Wolf, β -wolf, and δ -wolf are the leader wolves of the wolves, representing the optimal solution in evolution, and other wolves do not know the information of the optimal solution. Other wolves constantly adjust the search range and step size based on the position of the leader wolf, and the three best wolves formed after adjustment will become the leader wolves, so as to search for the best solution through continuous loop iteration. The algorithm simulates the encirclement, hunting, attack, and search of wolves.

(1) *Surround*. The positions of other wolves can be represented by the following formula:

$$\begin{aligned} D &= |C \cdot X_q(k) - X(k)|, \\ X(k+1) &= X_q(k) - A \cdot D, \\ A &= 2\zeta \cdot \tau_1 - \zeta, C = 2\tau_2. \end{aligned} \quad (13)$$

Type: X_q is the position of prey, which can be replaced by the mean value of α , β , and δ wolves; $X(k)$ is the K -generation individual wolf pack; τ_1, τ_2 are random numbers in the interval $[0,1]$; ζ is linearly decreasing from 2 to 0 based on the number of iterations.

(2) *Hunting*. In this step, it is assumed that α -wolf, β -wolf, and δ -wolf know the position of prey (optimal solution), and the algorithm uses these three positions to calculate the optimal solution. $D_\rho = |C_\rho \cdot X_\rho - X(k)|, \rho = \alpha, \beta, \delta$. At the same time, other wolves are approaching the optimal solution.

Algorithm: Grey wolf optimization algorithm

Inputs: objective function $f(x)$, search the upper and lower limits UB and LB of space, and set the number of individual wolves N and the maximum iteration number T ;

1: randomly initialize the population position $X_i (i=1, 2, \dots, N)$. Where $X_i=(x_{i1}, x_{i2}, \dots, x_{iD})$, and d represents the dimension, initialization ζ , a and $C, t=0$;

2: evaluate the fitness value of each search individual, where α -wolf = the best search individual, β -wolf = the second best search individual and δ -wolf = the third best search individual;

3: While $t < T$ do

4: for every individual in the wolf pack

5: update the current individual position and evaluate the individual according to formula (5. 24);

6: endfor

7: update α -wolf = optimal search individual, β -wolf = second optimal search individual and δ -wolf = third optimal search individual;

8: update ζ, A and C ;

9: $t=t+1$;

10: endwhile

Outputs: α wolf and its fitness value are the best points and optimal values.

FIGURE 5: The pseudocode of wolf optimization algorithm.

The positions of α -wolf, β -wolf, and δ -wolf can be represented by the following formula:

$$\begin{aligned} X'_\rho &= X_\rho - A_\rho \cdot D_\rho, \\ X(k+1) &= \frac{X'_\alpha + X'_\beta + X'_\delta}{3}. \end{aligned} \quad (14)$$

Type: D_ρ is the distance between α, β , and δ wolves and other individuals; X_ρ is the position of α, β , and δ wolves; $X(k)$ is the location of other individual wolves; X'_ρ is the position of other individual wolves moving towards α, β , and δ wolves.

(3) *Attack*. Through the continuous iteration of the algorithm, α, β , and δ wolves approach the optimal solution continuously, achieving the purpose of optimization. In the simulation of this process, ζ is linear decrease, from 2 to 0, and the range of a value is $[-\zeta, \zeta]$. When $|A| < 1$, the next position of the wolf pack will be closer to the position of the prey, and the algorithm will search locally. When $|A| > 1$, the distance between the next position of the wolf pack and the position of the prey becomes larger, and then the algorithm performs global search.

(4) *Search*. Individual wolves search for the best solution (prey) based on the position of α, β , and δ wolves. The algorithm is based on the relationship between $|A|$ value and 1 as the standard to divide global search and local search. In order to avoid the search falling into the local optimal solution, the grey wolf algorithm introduces the parameter C . The parameter C is a random value in the interval $[0,2]$ and decreases nonlinearly. $C > 1$ indicates that the random weight of the position of the wolf pack is significant, and $C < 1$ indicates that the random weight of the position of the wolf pack is small. Parameter c can make the algorithm search globally and can make the algorithm jump out of the local optimal solution.

See the detailed development process of the grey wolf algorithm and the pseudocode of the grey wolf algorithm as shown in Figure 5.

2.3. Selection of Optimization Methods

2.3.1. Solution Method. Based on the analysis of flood composition in Luanhe River Basin, taking the minimum square of the sum of reservoir discharge and interval flood discharge as the objective function, four optimization algorithms, namely, genetic algorithm (GA), particle swarm optimization (PSO), spider swarm optimization (SSO), and grey wolf optimization algorithm (GWO), are introduced into the optimal flood control operation to seek the minimum value of the objective function, and the results are compared and analyzed.

2.3.2. Introduction of Parameters. The computer is used for simulation experiment. The selected computer is Win10 and 64-bit operating system, the processor is 2.9 GHz, and the memory is 16 GB. MATLAB2018b is used for programming calculation. The number of samples is 50 and the maximum number of iterations is 500. The optimized parameters of the algorithm are obtained by experiment and debugging: the crossover probability of genetic algorithm is 0.8, and the mutation probability is 0.2; inertia weight of particle swarm optimization is ζ . It is a linearly decreasing value in the range of 0.9~0.4, $\delta_1 = \delta_2 = 2$. The probability factor pf_i of spider algorithm is 0~3 increasing value. The parameters of the grey wolf optimization algorithm have been explained above. This paper analyzes the optimization algorithm of flood dispatching in Panjiakou Reservoir and finds out the optimal method for optimal flood dispatching in Luanhe River Basin.

3. Results and Discussion

3.1. Optimization Results. Genetic algorithm, particle swarm optimization, spider swarm optimization, and grey wolf optimization algorithm are applied to the optimal flood control operation of Panjiakou Reservoir, and the above-mentioned models are applied to the optimal flood control operation of the basin once every 3 years, once every 5 years, once every 10 years, once every 20 years, and once every 50 years. The change of the result of each iteration of the optimization algorithm with the number of iterations is taken as the basic basis for analyzing the performance of the optimization algorithm.

From the analysis in Figure 6, it can be seen that, for the flood that occurs once every three years or once every five years, the grey wolf optimization algorithm is the best iteration of the four methods, and the grey wolf optimization algorithm, genetic algorithm, and particle swarm algorithm converge earlier, and the best convergence value of the grey wolf optimization algorithm is better than that of genetic algorithm and particle swarm algorithm. Spider swarm optimization algorithm converges late, but the optimal value of convergence is better than genetic algorithm and particle swarm optimization algorithm. For the once-in-ten-year flood, the best iteration of the four methods is the grey wolf optimization algorithm, which converges earlier than the spider swarm algorithm and the particle swarm algorithm,

and the best convergence value of the grey wolf optimization algorithm is better than that of the spider swarm algorithm and the particle swarm algorithm. Spider swarm algorithm converges later than spider swarm algorithm, and the objective function value of genetic algorithm changes little. According to the change trend in the figure, with the increase of iteration times, the spider swarm algorithm may continue to converge.

For the flood that occurs once every 20 years, the analysis shows that the change trend of genetic algorithm and particle swarm optimization is similar, particle swarm optimization is slightly better than genetic algorithm, and the convergence trend of genetic algorithm and particle swarm optimization is not obvious. The convergence performance of grey wolf optimization algorithm and spider swarm algorithm is better, and the convergence performance of grey wolf optimization algorithm is better than that of spider swarm algorithm. By analyzing the changing trend of the objective function, it can be seen that the spider swarm algorithm may still converge with the increase of iteration times.

For the once-in-50-year flood, the analysis shows that the convergence performance of genetic algorithm and particle swarm optimization is not obvious and basically maintains the initial value. The convergence performance of spider swarm algorithm and grey wolf optimization algorithm is better than that of spider swarm algorithm. Spider swarm optimization is basically stable after convergence reaches a certain objective function value. The grey wolf optimization algorithm converges faster and better. Based on the changing trend, it is known that the grey wolf optimization algorithm still has the possibility of continuous iterative convergence.

3.2. Scheduling Results. The results of reservoir operation are evaluated by the sum of squares of outflow and peak clipping rate, and the results of flood control optimal operation are analyzed and calculated. The results are shown in Table 1. As can be seen from Table 1, for the analysis of the sum of squares of the outflow, the spider swarm algorithm and the grey wolf optimization algorithm of the inflow flood once every 3 years, 5 years, 10 years, and 20 years are superior to the genetic algorithm and the particle swarm algorithm. The grey wolf optimization algorithm is the best. The 50-year return spider swarm algorithm and particle swarm algorithm have similar results, the worst is genetic algorithm, and the best is grey wolf optimization algorithm. According to the analysis of peak clipping rate of the inflow flood with the frequency of 5 years, 10 years, 20 years, and 50 years, the order of peak clipping performance from high to low is grey wolf optimization algorithm > spider swarm algorithm > particle swarm algorithm > genetic algorithm. For the flood that occurs once every three years, the order of peak clipping performance from high to low is grey wolf optimization algorithm > particle swarm algorithm > spider swarm algorithm > genetic algorithm.

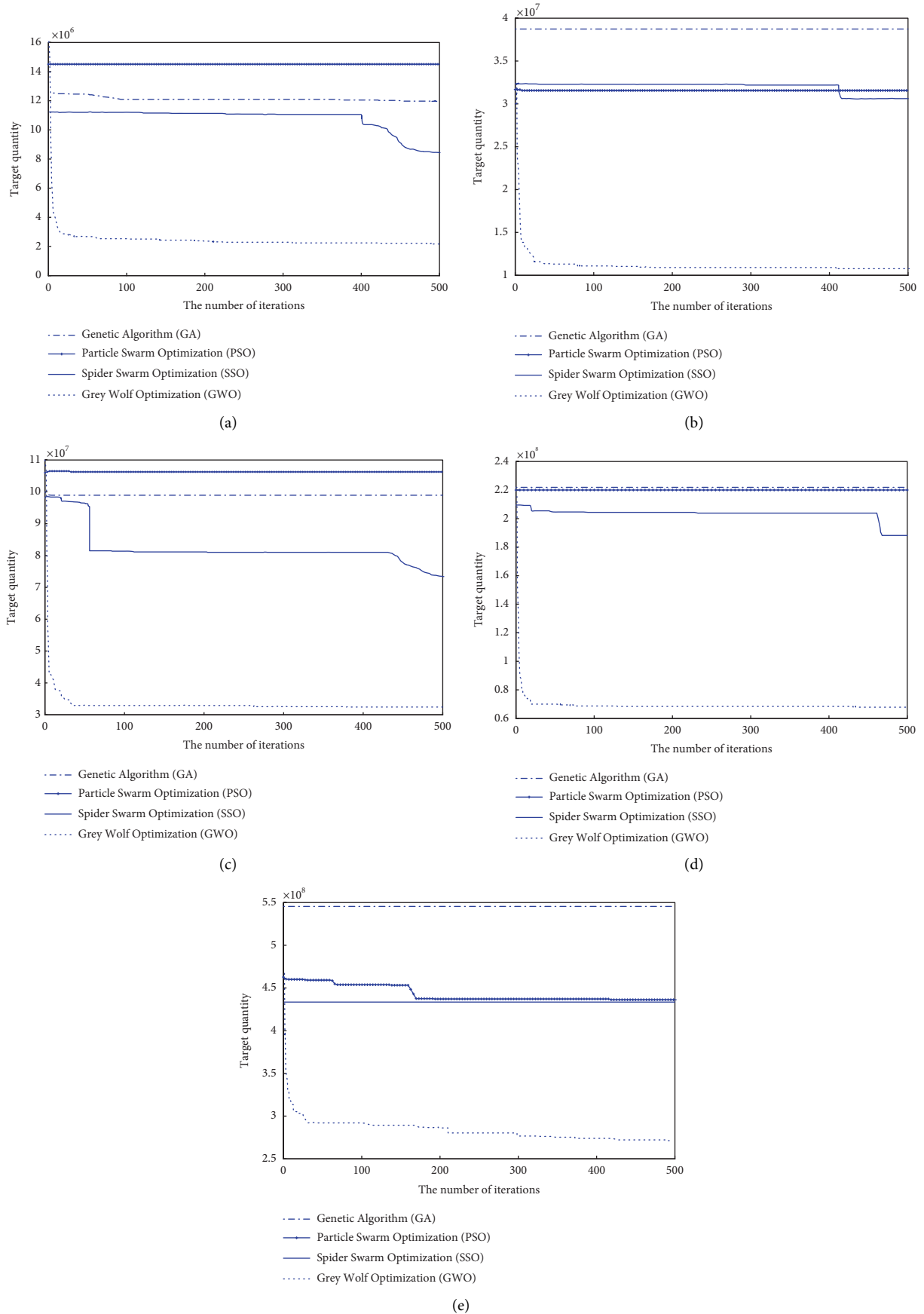


FIGURE 6: Iterative process of Panjiakou Reservoir operation with 3-year return period (a), 5-year return period (b), 10-year return period (c), 20-year return period (d), and 50-year return period (e).

TABLE 1: Statistical table of inflow flood regulation results.

Eigenvalue	Way	Once in three years	Once in five years	Once in 10 years	Once in 20 years	Once in 50 years
<i>Sum of squares of flow</i>	Genetic algorithm	5.47×10^6	3.88×10^7	5.91×10^7	8.22×10^7	3.90×10^8
	Particle swarm optimization algorithm	4.25×10^6	3.15×10^7	4.06×10^7	8.20×10^7	3.34×10^8
	Spider swarm algorithm	4.42×10^6	3.06×10^7	4.36×10^7	7.88×10^7	3.37×10^8
	Grey wolf optimization algorithm	3.96×10^6	1.08×10^7	3.26×10^7	6.81×10^7	2.75×10^8
<i>Peak clipping rate (%)</i>	Genetic algorithm	59.65	65.62	59.73	63.52	64.09
	Particle swarm optimization algorithm	67.82	61.59	61.32	64.55	65.92
	Spider swarm algorithm	61.98	69.48	64.69	65.69	66.23
	Grey wolf optimization algorithm	70	70	70	70	69.90

4. Conclusions

Genetic algorithm (GA), particle swarm optimization (PSO), spider swarm optimization (PSO), and grey wolf optimization (GWO) are applied to the optimal operation of reservoir flood control.

- (1) Optimization ability: the analysis shows that genetic algorithm and particle swarm optimization algorithm have obvious premature convergence characteristics and poor iterative optimization performance. The spider swarm algorithm has strong optimization ability, and its optimization ability has strong correlation with iteration times and optimization time. Grey wolf optimization algorithm has strong optimization ability, good convergence effect, and stable results. Analysis shows that grey wolf optimization algorithm is better than spider swarm algorithm for long-term flood control operation. There are two types of spiders, male and female, and grey wolf optimization algorithm α , β , and δ wolves and other types of wolves are the main particles for optimization. It shows that dividing the artificial intelligence population into different types of elite particles can expand the search range and search ability from different angles, so as to effectively improve the optimization ability of the algorithm. The simulation formula of male and female spiders of spider swarm algorithm is complex, while the simulation of grey wolf optimization algorithm is simple and easy to understand and operate, which shows that grey wolf optimization algorithm has a good model structure. It can make full use of the results of the three wolf groups to optimize.
- (2) Scheduling results: from the analysis of the square sum of the outflow flow and the peak clipping rate, the results of genetic algorithm and particle swarm optimization algorithm are similar, and the results of spider swarm algorithm and grey wolf optimization algorithm are similar and slightly better than the first two. Combined with the optimization ability, it can be expressed as follows: the algorithm with better

optimization performance has strong peak clipping ability and has good peak clipping ability for each period in the iterative calculation process. The peak clipping rate of grey wolf optimization algorithm is more stable than spider swarm algorithm.

Based on the above analysis, it can be seen that grey wolf optimization algorithm has better optimization performance, and the Panjiakou Reservoir group data series is longer. Grey wolf optimization algorithm has a wider search range and shorter calculation time in solving the long sequence problem. Therefore, grey wolf optimization algorithm can be applied to solve the flood control operation optimization model of Panjiakou Reservoir Group.

Data Availability

The figures and tables used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to show sincere thanks to those technicians who have contributed to this research. This research was supported by Key Laboratory of Lower Yellow River Channel and Estuary Regulation Foundation (LYRCER202103) and also by Special Basic Research Fund for Central Public Research Institutes (HKY-JBYW-2020-15).

References

- [1] J. Li, M. Zhang, E. Jiang et al., "Influence of floodplain flooding on channel siltation adjustment under the effect of vegetation on a meandering riverine beach," *Water*, vol. 13, no. 10, p. 1402, 2021.
- [2] Y. Liu, Y. Wang, and E. Jiang, "Stability index for the planview morphology of alluvial rivers and a case study of the Lower

- Yellow River,” *Geomorphology*, vol. 389, Article ID 107853, 2021.
- [3] B. Xue, J. Li, Y. Zhang, and L. Xu, “Experimental simulation of river environment transformation and development,” *Freshwater Environmental Bulletin*, vol. 30, no. 9, Article ID 10675, 2021.
- [4] L. Yu, G. Shenglian, and Z. Yanlai, “Flood control optimal operation of cascade reservoirs considering stochastic process of inflow flood,” *Journal of Sichuan University Engineering Science Edition*, vol. 4, no. 6, pp. 13–20, 2012.
- [5] Z. Shuanghu, H. Qiang, and W. Hongshou, “An improved particle swarm optimization algorithm for optimal reservoir operation of hydropower stations,” *Journal of Hydropower*, vol. 26, no. 1, pp. 1–5, 2007.
- [6] J. W. Labadie, “Optimal operation of multireservoir systems: state-of-the-art review,” *Journal of Water Resources Planning and Management*, vol. 130, no. 2, pp. 93–111, 2004.
- [7] S. Gupta and K. Deep, “A novel hybrid sine cosine algorithm for global optimization and its application to train multilayer perceptrons,” *Applied Intelligence*, vol. 50, no. 4, pp. 993–1026, 2020.
- [8] K. Zhou, “Grey nonlinear water environment management model based on genetic algorithm and its application,” *People’s president*, vol. 50, no. 5, pp. 20–40, 2019.
- [9] L. Cong, F. Wei, and H. Sheng, “A survey of research and application of grey wolf optimization algorithm,” *Science Technology and Engineering*, vol. 20, no. 9, pp. 3378–3386, 2020.
- [10] Z. Wenhuan, L. Rongbo, and Z. Li, “Risk flood control operation analysis of reservoirs in the Yangtze River Basin,” *People’s Yangtze River*, vol. 51, no. 12, pp. 135–178, 2020.
- [11] X. Zhang, R. Srinivasan, and M. V. Liew, “On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site calibration of the SWAT model,” *Hydrological Processes*, vol. 24, no. 8, pp. 955–969, 2010.
- [12] J. Li and P. Feng, “Runoff variations in the Luanhe River basin during 1956–2002,” *Journal of Geographical Sciences*, vol. 17, no. 3, pp. 339–350, 2007.
- [13] P. Ren, J. Li, P. Feng, Y. Guo, and Q. Ma, “Evaluation of multiple satellite precipitation products and their use in hydrological modelling over the Luanhe River basin, China,” *Water*, vol. 10, no. 6, p. 677, 2018.
- [14] J. Liu, Q. Chen, and Y. Li, “Ecological risk assessment of water environment for Luanhe River Basin based on relative risk model,” *Ecotoxicology*, vol. 19, no. 8, pp. 1400–1415, 2010.
- [15] O. I. Unver and L. W. Mays, “Model for real-time optimal flood control operation of a reservoir system,” *Water Resources Management*, vol. 4, no. 1, pp. 21–46, 1990.
- [16] G. Fu, “A fuzzy optimization method for multicriteria decision making: an application to reservoir flood control operation,” *Expert Systems with Applications*, vol. 34, no. 1, pp. 145–149, 2008.
- [17] Y. He, Q. Xu, S. Yang, and L. Liao, “Reservoir flood control operation based on chaotic particle swarm optimization algorithm,” *Applied Mathematical Modelling*, vol. 38, no. 17–18, pp. 4480–4492, 2014.
- [18] D. Che and L. W. Mays, “Development of an optimization/simulation model for real-time flood-control operation of river-reservoirs systems,” *Water Resources Management*, vol. 29, no. 11, pp. 3987–4005, 2015.
- [19] H. Qin, J. Zhou, Y. Lu, Y. Li, and Y. Zhang, “Multi-objective cultured differential evolution for generating optimal trade-offs in reservoir flood control operation,” *Water Resources Management*, vol. 24, no. 11, pp. 2611–2632, 2010.
- [20] S. G. Gino Sophia, V. Ceronmani Sharmila, S. Suchitra, T. Sudalai Muthu, and B. Pavithra, “Water management using genetic algorithm-based machine learning,” *Soft Computing*, vol. 24, no. 22, Article ID 17153, 2020.
- [21] D. Ma, H. Duan, W. Li, J. Zhang, W. Liu, and Z. Zhou, “Prediction of water inflow from fault by particle swarm optimization-based modified grey models,” *Environmental Science and Pollution Research*, vol. 27, no. 33, Article ID 42051, 2020.
- [22] A. M. Ibrahim, M. A. Tawhid, and R. K. Ward, “A binary water wave optimization for feature selection,” *International Journal of Approximate Reasoning*, vol. 120, pp. 74–91, 2020.

Research Article

Prediction of Response to Radiotherapy by Characterizing the Transcriptomic Features in Clinical Tumor Samples across 15 Cancer Types

Yu Xu,¹ Chao Tang,² Yan Wu,² Ling Luo,² Ying Wang,² Yongzhong Wu ,² and Xiaolong Shi ²

¹College of Bioengineering, Chongqing University, Chongqing, China

²Radiation and Cancer Biology Laboratory, Radiation Oncology Center, Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Institute and Chongqing Cancer Hospital, Chongqing University Cancer Hospital and Chongqing Cancer, Chongqing 400030, China

Correspondence should be addressed to Yongzhong Wu; cqmdwyz@163.com and Xiaolong Shi; xshi.bear@cqu.edu.cn

Received 13 March 2022; Revised 12 April 2022; Accepted 13 April 2022; Published 9 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yu 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.

Purpose. Radiotherapy (RT) is one of the major cancer treatments. However, the responses to RT vary among individual patients, partly due to the differences of the status of gene expression and mutation in tumors of patients. Identification of patients who will benefit from RT will improve the efficacy of RT. However, only a few clinical biomarkers were currently used to predict RT response. Our aim is to obtain gene signatures that can be used to predict RT response by analyzing the transcriptome differences between RT responder and nonresponder groups. **Materials and Methods.** We obtained transcriptome data of 1664 patients treated with RT from the TCGA database across 15 cancer types. First, the genes with a significant difference between RT responder (R group) and nonresponder groups (PD group) were identified, and the top 100 genes were used to build the gene signatures. Then, we developed the predictive model based on binary logistic regression to predict patient response to RT. **Results.** We identified a series of differentially expressed genes between the two groups, which are involved in cell proliferation, migration, invasion, EMT, and DNA damage repair pathway. Among them, MDC1, UCP2, and RBM45 have been demonstrated to be involved in DNA damage repair and radiosensitivity. Our analysis revealed that the predictive model was highly specific for distinguishing the R and PD patients in different cancer types with an area under the curve (AUC) ranging from 0.772 to 0.972. It also provided a more accurate prediction than that from a single-gene signature for the overall survival (OS) of patients. **Conclusion.** The predictive model has a potential clinical application as a biomarker to help physicians create optimal treatment plans. Furthermore, some of the genes identified here may be directly involved in radioresistance, providing clues for further studies on the mechanism of radioresistance.

1. Introduction

Cancer is the leading cause of morbidity and mortality in the world, regardless of the level of human development, which is accounting for over 9.9 million deaths worldwide annually [1, 2]. Radiation therapy (RT) has played a major role in cancer therapeutics, with approximately 50% of cancer patients using RT [3]. RT mainly induces apoptosis of cancer cells by causing DNA double-strand breaks (DSBs) [4]. RT is

often curative, particularly in some head and neck cancer, prostate cancer, and cervical cancer. However, the inherent radioresistance of tumor cells and the acquired radioresistance can reduce the therapeutic effect and ultimately lead to poor outcomes of patients such as tumor recurrence, metastasis, and patient death [5]. In the era of precision medicine, there is a growing need for precision radiotherapy. This requires the plans of individualized treatment by considering patients' multiple information, including the

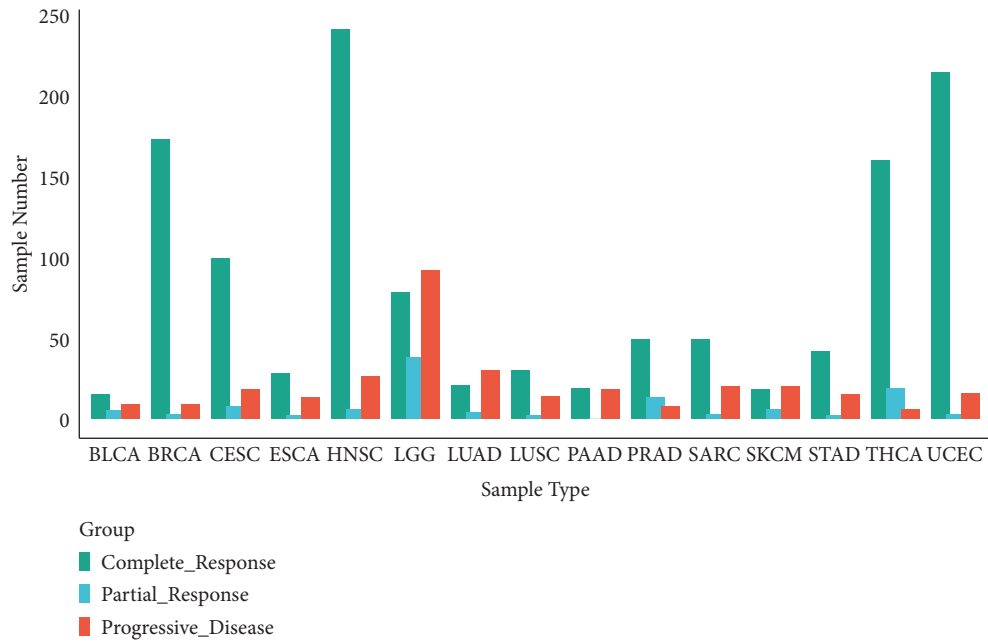


FIGURE 1: Statistical histogram of cancer types of samples and grouping.

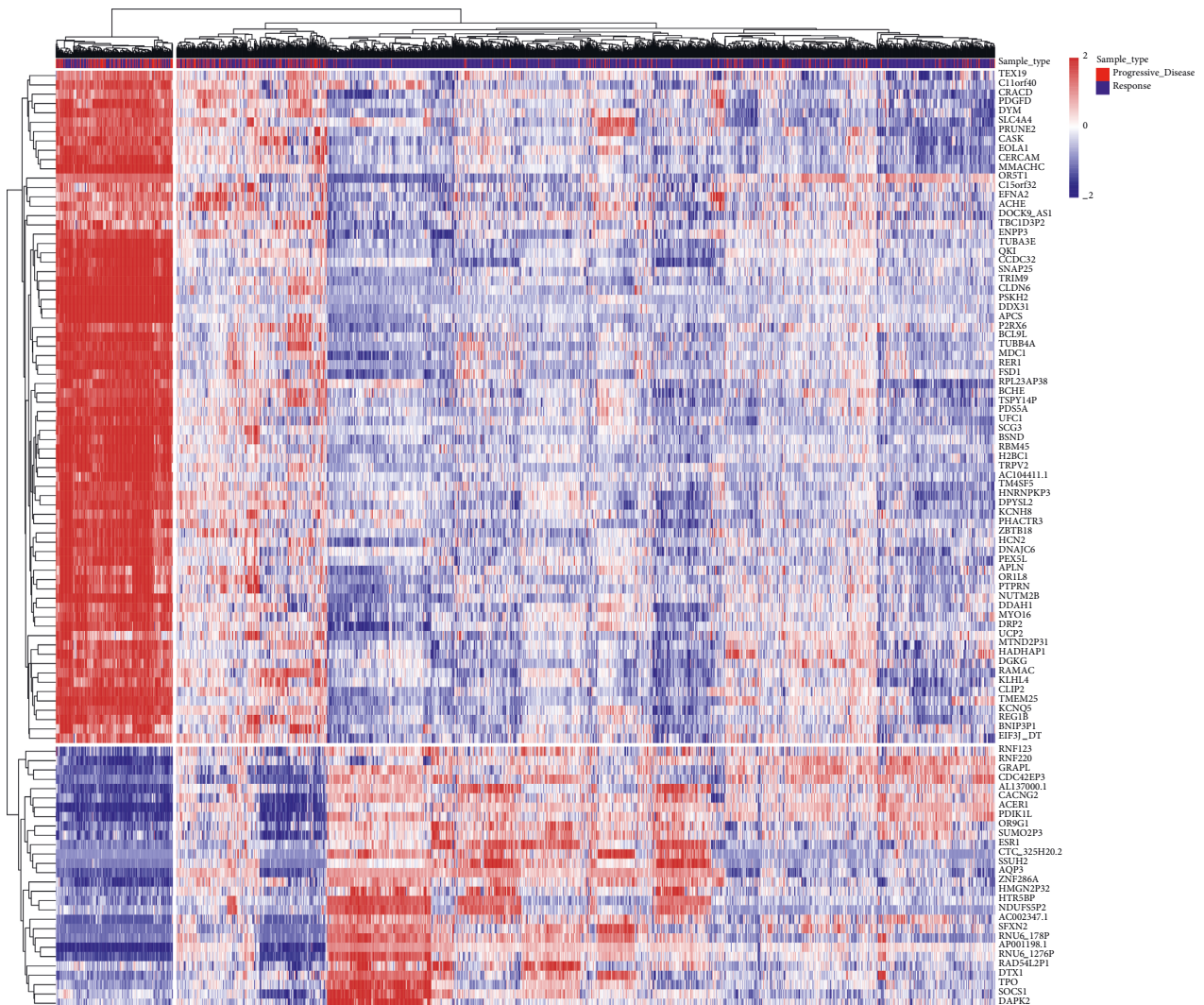


FIGURE 2: Heatmap of the top 100 differentially genes for all cancer species analyzed. Both samples and genes were clustered with average linkage.

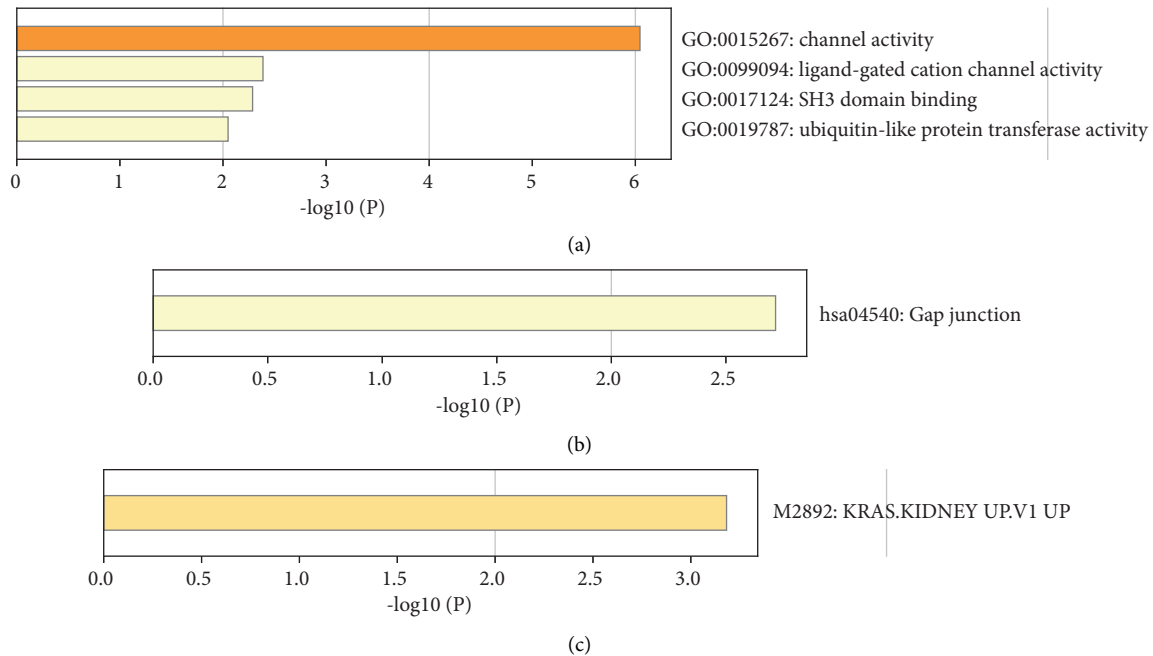


FIGURE 3: Histogram of pathway enrichment of top 100 differential genes.

status of gene mutation and profile of gene expression, in order to achieve the best therapeutic efficacy. Radioresistance is partly due to the differences of the status of gene expression and mutation in tumors of patients. Thus, identification of patients who will benefit from RT through analyzing gene mutation or gene expression will improve the efficacy of RT.

The previous studies have demonstrated that the mutations of several genes influenced the radiosensitivity. Ataxia-telangiectasia mutated (ATM) is a central regulator of DNA damage response [6]. Mutation and inactivation of ATM can lead to increased gene instability and impaired repair of DNA double-strand breaks. Tumors harboring ATM mutation are extremely sensitive to radiation. Mutations of ATM have been found in several tumor types. Targeted next-generation sequencing has revealed an 8% incidence of ATM mutations in prostate cancer [7]. ATM is also one of the commonly mutated genes in mantle cell lymphoma (MCL) and lung adenocarcinoma [7, 8]. Recent experimental evidence demonstrated that the mutational status of ATM can be used as a biomarker for radiotherapy [9]. The correlation between KEAP1/NRF2 mutation status and radioresistance has been investigated. Keap1-Nrf2 is a master regulator of cellular response to oxidative and radiation stress [10]. The KEAP1-NRF2 pathway is involved in protection of cells from oxidative and toxic stresses. Recent studies indicated that the activating mutation of the KEAP1-NRF2 pathway induces radioresistance, and KEAP1/NRF2 mutation status is a strong predictor of RT outcome in patients with NSCLC [11].

Although the mutations of these genes could serve as predictive biomarkers for personalization of therapeutic strategies, their widespread use would be limited due to the limited number of ATM or KEAP1-NRF2 mutations carriers

among cancer patients. The previous studies revealed that the difference in the status of gene expression (including lncRNA and miRNA) and DNA methylation on tumor suppressor gene may also contribute to resistance to RT. For example, radioresistance in NSCLC has been associated with overexpression of antioxidant enzymes such as Mn-superoxide dismutase (Mn-SOD) [12]. Several studies revealed many miRNAs, including miR-95 [13], miR-221, miR-222 [14], and miR-106b [15], enhanced radioresistance in cancer cells, while miR-30a [16], miR-16 [17], miR-449 [18], miR-17 [19], and miR-100 [20] enhance the radiosensitivity. The role of epigenetic modifications, especially DNA methylation, has been explored in radioresistance in malignant tumors. The changes in radiosensitivity caused by DNA methylation in the promoter region of genes associated with cell damage repair, cell proliferation, and cell cycle have been demonstrated [21]. For example, Liu et al. [22] found the promoter region of the ERCC1 gene in two radiosensitive cell lines was hypermethylated, while it was hypomethylated in the other two radioresistant cell lines. These studies provided potential biomarkers for stratification of RT patients.

In previous studies, cancer cell lines usually were used as a model for radioresistance study. We speculated that radioresistance probably involves a complex transcriptional coexpression network within tumor cells. In tumor tissue, the factors involved in the response to radiation might be more complex than in cell lines. Therefore, biomarkers obtained from the analysis of actual clinical tumor samples may be closer to clinical application. The development of high-throughput sequencing methods for mRNA (RNA-Seq) has provided a very powerful tool to analyze the transcriptome of tumors, and a huge amount of data are available [23]. In this study, we performed a large-scale analysis of transcriptomic data collected from The Cancer

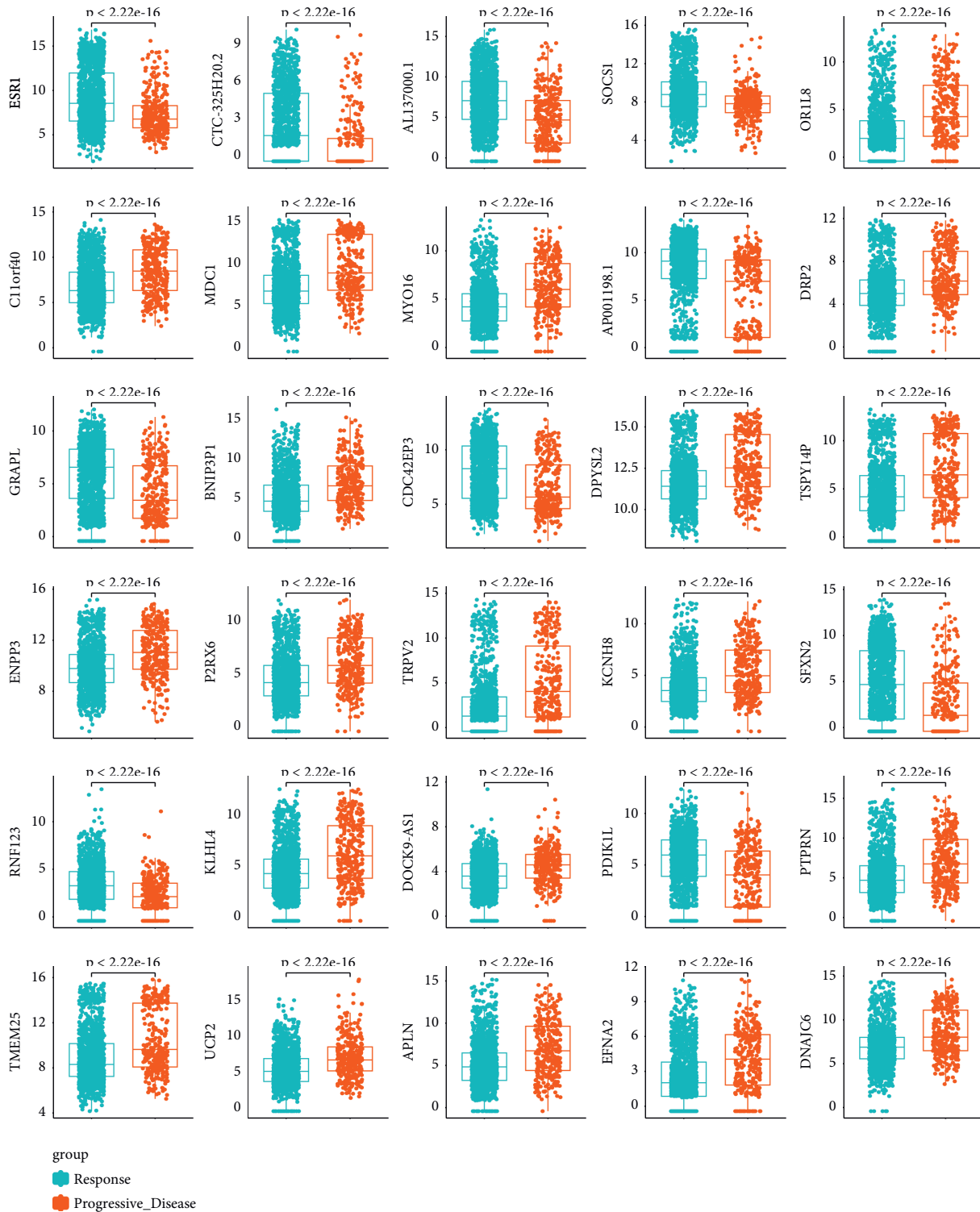


FIGURE 4: mRNA levels of differentially expressed genes between response and progressive disease tumors. The distribution of gene expression values in R or PD samples was drawn through the boxplot, and the p value is marked (t -test).

Genome Atlas (TCGA) from those who were treated with RT across 15 different cancer types. The gene signatures for predicting RT response were obtained by analyzing the transcriptome differences between RT responder and non-responder groups, and the performance of gene signatures was estimated in this study.

2. Materials and Methods

2.1. Data Acquisition. The data for this study were obtained from the TCGA website (<https://portal.gdc.cancer.gov/>). This study downloaded 15 types of cancer clinical data and gene expression data. Fifteen types of cancers include

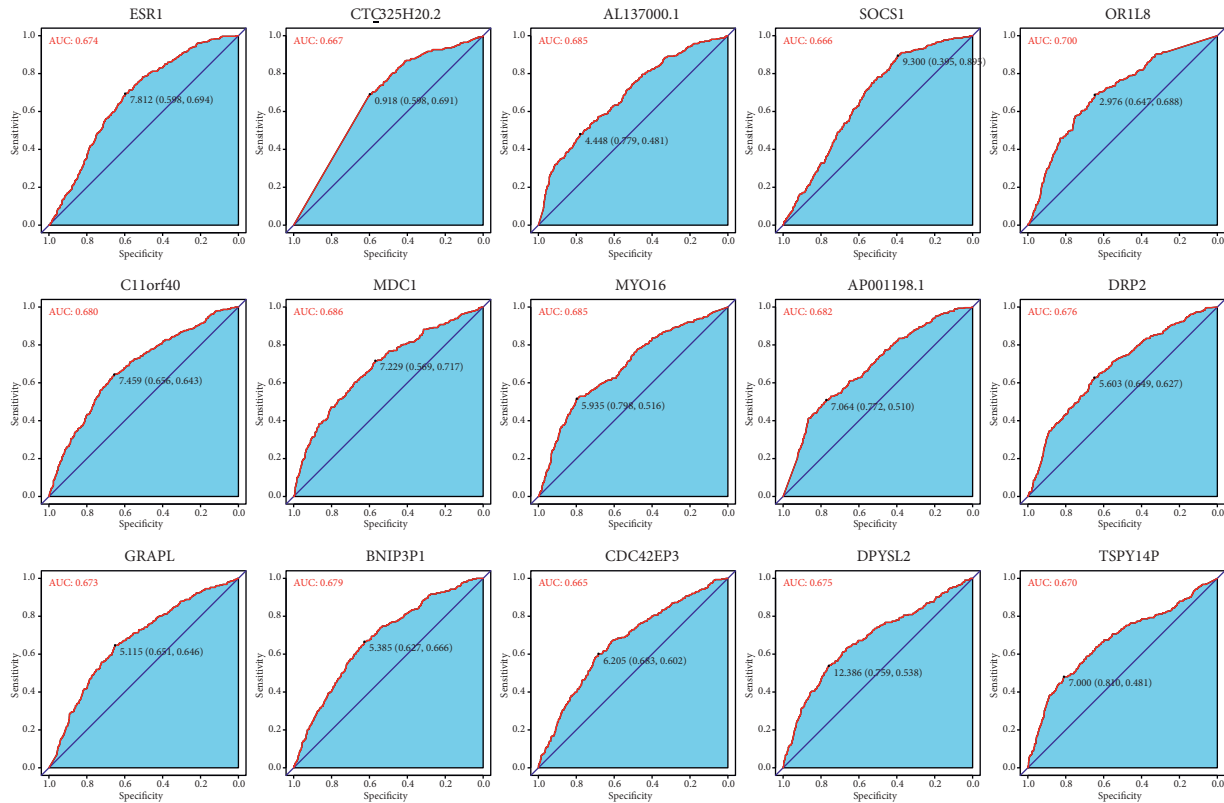


FIGURE 5: ROC curve of differentially expressed genes between response tumors and progressive disease tumors.

bladder urothelial carcinoma (BLCA), invasive breast cancer (BRCA), cervical squamous cell carcinoma (CESC), esophageal cancer (ESCA), head and neck squamous cell carcinoma (HNSC), low-grade glioma (LGG), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), pancreatic cancer (PAAD), prostate adenocarcinoma (PRAD), sarcoma (SARC), cutaneous melanoma (SKCM), gastric adenocarcinoma (STAD), thyroid carcinoma (THCA), and endometrial carcinoma (UCEC). First, this study extracted the clinical treatment information and overall survival information of patients. Next, tumor tissue samples from those patients treated with radiotherapy were selected according to clinical data, and then the samples were divided into two categories according to RECIST response results: response group (including complete response and partial response) and disease progression group. Finally, 1664 tumor samples were used for analysis in this study, including 1350 response (R) samples and 314 progressive disease (PD) samples.

2.2. Expression Analysis and Differential Expression Gene Identification. Data of the mRNA profile of samples selected from the TCGA dataset were analyzed using the DESeq2 package [24] in R language. In the gene expression data, the genes will not be analyzed further if the read counts of these genes were less than 10 in 80% of samples. In this study, samples were divided into response (R) and progressive disease (PD) samples. The differentially expressed genes

(DEGs) between R and PD samples were identified at the criteria of $p < 0.05$ and $|\log FC| > 1$. Wald test was used to calculate the p value. The t -test was also used to calculate p value in this study. The top 100 genes with the significant difference were obtained for follow-up analysis. For individual tumor types, the tumor type-specific gene signature was obtained from the top 100 genes according to the p value of each gene expression difference in individual tumor types.

2.3. Binary Logistic Regression Analysis. Logistic regression analysis was used to establish a prediction model based on risk factors and used one or more explanatory variables to predict a class of response variables [25]. The standardized expression value of the top 100 genes was used as the variable; the R sample was used as the reference group, and the value was set to 1; the value of progressive sample was set to 0. The RT response was predicted by logistic regression modeling and quantified using a prediction index. The prediction index is the regression coefficient obtained by multiplying the expression level of the selected gene with the logistic regression model (β). The prediction index for each patient is as follows: predictive index = $\text{Expr gene1} * \beta_{\text{gene1}} + \text{Expr gene2} * \beta_{\text{gene2}} + \text{Expr gene3} * \beta_{\text{gene3}} + \dots + \text{Expr geneN} * \beta_{\text{geneN}} + \text{intercept}$.

If the prediction index is less than the threshold (e.g., < 0.5), the sample is more likely to be a progressive disease, and if the prediction index is more than the threshold, the sample is more likely to be radiosensitive.

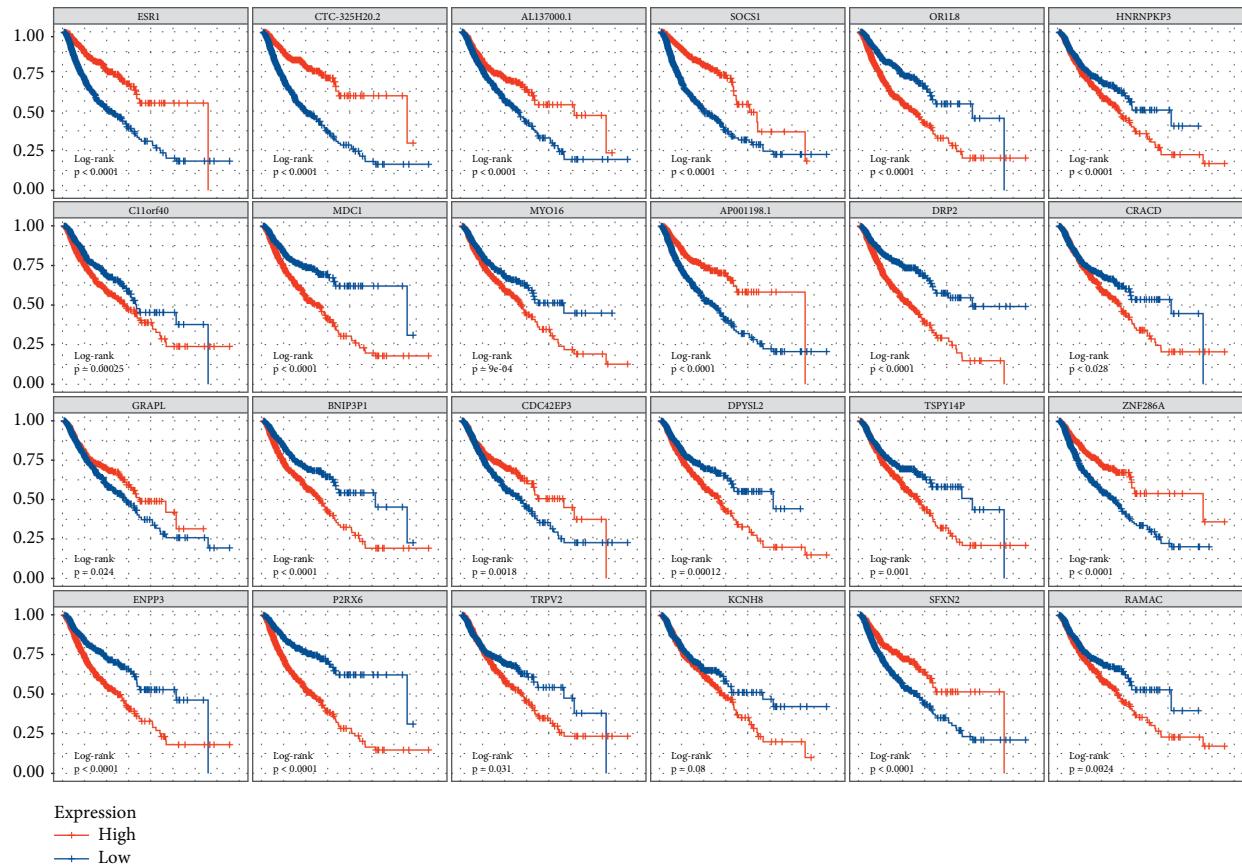


FIGURE 6: Overall survival curve of differential expression between reactive tumor and progressive disease tumor. The abscissa of the survival curve is the observation time and the ordinate is the survival rate. The median expression level of each gene was taken as the threshold, the high expression level group was higher than the median, and the low expression level group was lower than the median. The log-rank test was used to test the statistical significance of the two groups of data.

2.4. K-Fold Cross Validation. In the logistic regression model, the data set was divided into a training set and a validation set. In this study, the tumor expression level data set was divided into k subsets ($k = 10$) by using k -fold cross validation, and the method was repeated ten times. In each repetition, one subset was randomly selected as the validation set, and the remaining $k - 1$ subset was used as the training set. The training set data were used for establishing a logistic regression model, and then validation set data were brought into the model for evaluation. In each repetition, each evaluation score was retained, while the model was discarded. After ten repetitions, the model evaluation score was used to summarize the ability of logistic regression model. AUC (area under the curve) was used to evaluate a degree or measure of separability.

2.5. Visualization of Differential Expression Gene. The graphics of differentially expressed genes were drawn on the basis of R. Survminer package was used to analyze and visualize the survival curve of differentially expressed genes, and log-rank test was used to calculate the p value. The overall survival curve was divided into high expression group and low expression group according to the gene expression level. A box plot was drawn using the ggpubr

package to show the distribution of DEG in R and PD samples, and the p value was calculated using the t -test. The ROC curve of DEGs was drawn by using the pROC software package, and the best threshold to distinguish between R and PD samples was marked. Heatmap of DEGs was drawn by using pheatmap package. In the regression model, the samples were sorted from small to large according to the prediction index, and figures of prediction index were drawn by using the ggplot2 package. The best threshold point of AUC was used as a dividing line between R and PD samples.

2.6. Enrichment Pathway Analysis. Enrichment analysis of the top 100 differential genes was performed by the online tool Metascape (<https://metascape.org/>).

3. Result

3.1. Patient Characteristics. In order to describe the transcriptome characteristics of human cancer response to radiotherapy, the transcriptome of 1664 TCGA clinical samples from 15 cancer types was analyzed, including bladder urothelial carcinoma (BLCA, $n = 29$), breast invasive carcinoma (BRCA, $n = 185$), cervical squamous cell carcinoma (CESC, $n = 125$), esophageal carcinoma (ESCA,

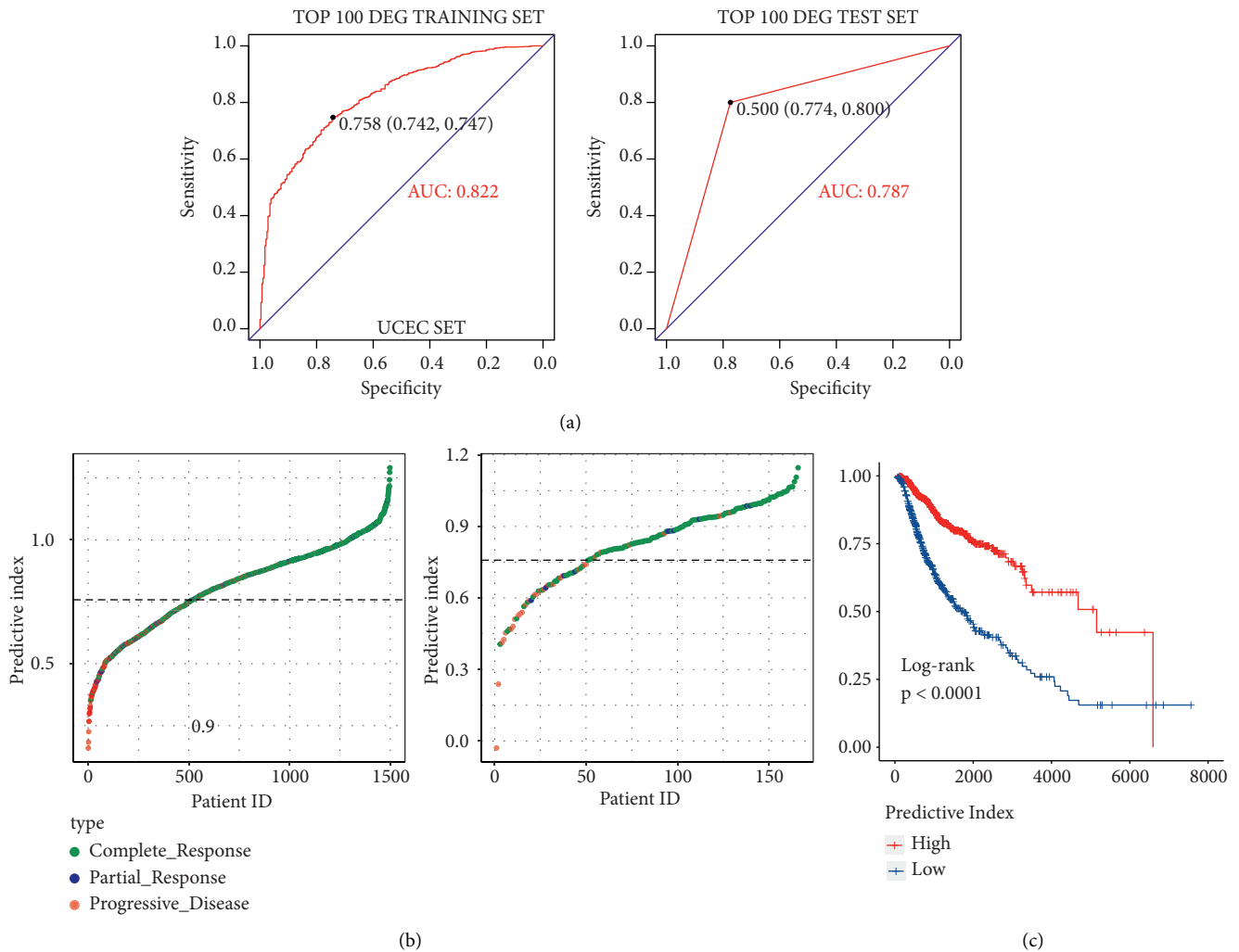


FIGURE 7: ROC curve, predictor chart, and overall survival rate chart of reactive tumor and progressive disease tumor after logistic regression. (a) ROC curve. (b) The ordinate is sorted from small to large according to the prediction index, and the threshold is 0.758. (c) Overall survival curve.

$n = 43$), head and neck squamous cell carcinoma (HNSC, $n = 273$), low-grade glioma (LGG, $n = 208$), lung adenocarcinoma (LUAD, $n = 55$), LUSC ($n = 46$), pancreatic cancer (PAAD, $n = 37$), prostate adenocarcinoma (PRAD, $n = 70$), sarcoma (SARC, $n = 72$), cutaneous melanoma (SKCM, $n = 59$), gastric adenocarcinoma (STAD, $n = 59$), thyroid carcinoma (THCA, $n = 44$), and endometrial carcinoma (UCEC, $n = 233$). The response group (R) was defined as patients with a relatively good response after RT (such as complete or partial response, $n = 1350$). The progressive disease group (PD) was defined as patients with imaging progressive disease after RT ($n = 314$). The cancer types and groups of samples were shown in Figure 1.

3.2. Identification of Differentially Expressed Genes in R Group and PD Group. First, according to the DEGs selection criteria $|\log_2 fc| > 1$, $p < 0.05$, the differentially expressed genes between R and PD were identified in 15 analyzed cancer types. Then, the t -test is carried out using these differential

genes to identify genes associated with R or PD (cutoff threshold $p < 0.05$). The heatmap was drawn using genes with the top 100 differential genes (Figure 2). Among the top 100 differential genes, the high expression of 72 genes and the low expression of 28 genes were highly correlated with PD samples. The functional annotation for these genes (Table S1) provides insight into the underlying biological mechanism leading to RT resistance. Genes involved in migration, cell proliferation, cell invasion, tumor metastasis, and EMT were significantly upregulated in the PD group (e.g., MDC1, UCP2, RBM45, BCL9L, P2RX6, RER1, EFNA2, CASK, CERCAM, and PTPRN). Interestingly, MDC1 has been reported as a key regulator of the DNA damage response in higher eukaryotes [26], and UCP2 and RBM45 have been implicated in RT resistance [27, 28]. Among the top 100 genes, there was a class of genes related to a ubiquitination proteasome hydrolysis system, such as RBM45, TRIM9, PTPRN, RNF123, RNF220, and DTX1. Ubiquitination is an important means of regulating target genes at the protein level. Several noncoding RNAs were also

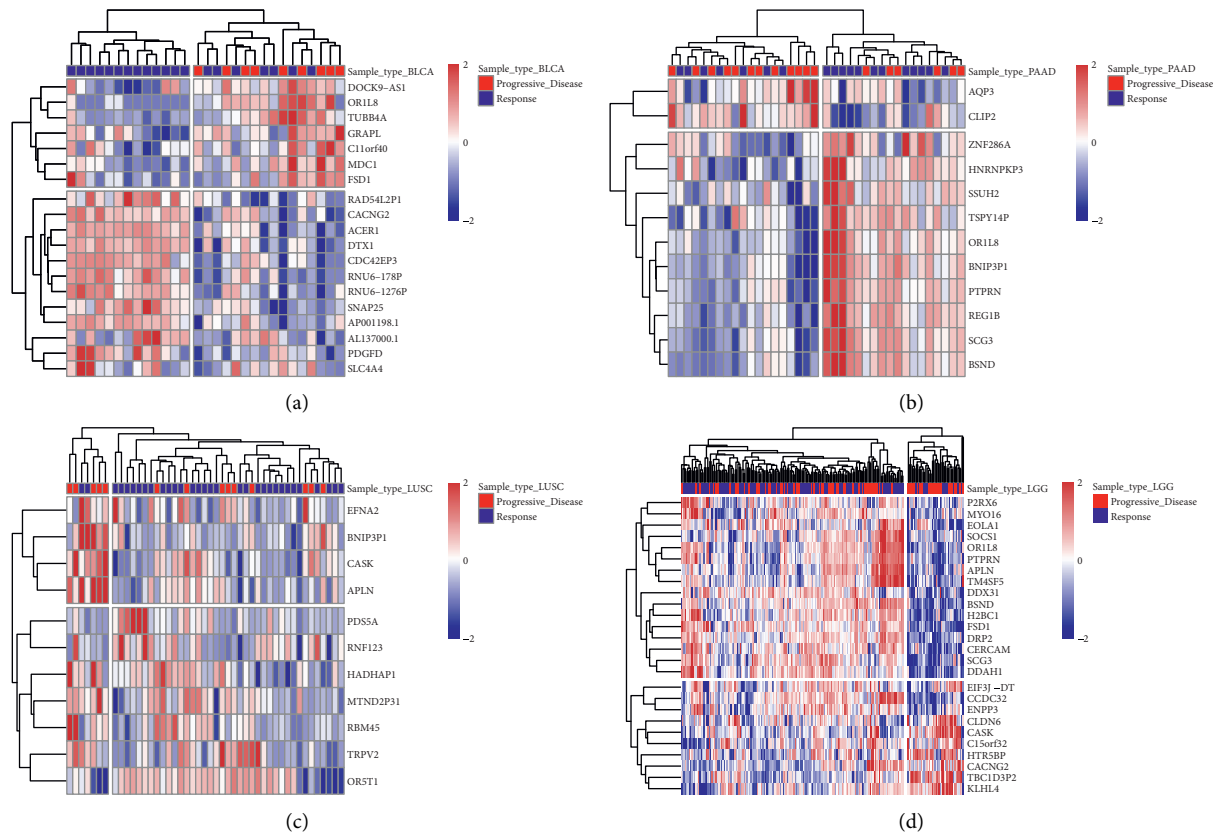


FIGURE 8: Heatmap of differentially expressed genes in different tumor types.

found in this study, including AC104411.1, CTC-325H20.2, AP001198.1, C15ORF32, C11ORF40, MTND2P31, RNU6-1276P, AC002347.1, RNU6-178P, HMG2P32, and EIF3J-DT.

GO analysis of the top 100 differential genes including biological process (BP), molecular function (MF), and oncogenic signatures (OncSig) analysis was conducted. In terms of MF, these genes are significantly enriched in channel activity, ligand-gated cation channel activity, SH3 domain binding, and ubiquitin-like protein transferase activity. BP analysis showed that the gap junction was significantly enriched. OncSig analysis showed that four genes (TRIM9, SCG3, SNAP25, and TUBB4A) were related to overexpression of KRAS. Our analysis revealed that the four genes tend to be highly expressed in PD samples. Their roles in cancer have been studied. For example, TRIM9 [29] and SCG3 [30] have been reported to promote the proliferation of cancer cell, and SNAP25 and TUBB4A were identified as potential prognostic biomarkers for prostate cancer [31] or lung adenocarcinoma [32] (Figure 3)

3.3. Assessing the Discriminative Power with a Single-Gene Signature. In this study, we identified the differentially expressed genes. We visualized the differential genes using the boxplot to estimate the discriminative power for R and PD samples with the single-gene signature (Figure 4). These genes show that there is a statistically significant differential

expression between R and PD samples. We further used the receiver operating characteristic (ROC) curve to evaluate the specificity of single-gene signature. As shown in Figure 5, the AUC value of single-gene discrimination model was between 0.615 and 0.7, and OR1L8 gene has the highest specificity (AUC = 0.700), implying that the discriminative power of single-gene signature is not very strong. However, the expression patterns of many genes may reflect the underlying molecular mechanisms of tumor radioresistance. For example, the high expression of MDC1 will increase the ability of DNA damage repair and eventually lead to an increase in radiotherapy resistance. Therefore, we believe that the expression patterns of some of these genes still have clinical predictive power.

3.4. Evaluation of Survival Rate by the Single-Gene Signature.

In order to investigate whether the single-gene signature has the ability to predict the survival rate of patients, we analyzed the overall survival data of the patients. The overall survival curve is shown in Figure 6. Statistically significant correlations were observed between the survival time and the expression in many chosen genes. For example, MDC1 ($p < 0.0001$), EFNA2 ($p < 0.0001$), BCL9L ($p < 0.0001$), RER1 ($p < 0.0001$), and P2RX6 ($p < 0.0001$) tended to be highly expressed in PD samples, and the overall survival rate of the high expression group was significantly lower than that of the low expression group. Conversely, ESR1 ($p < 0.0001$),

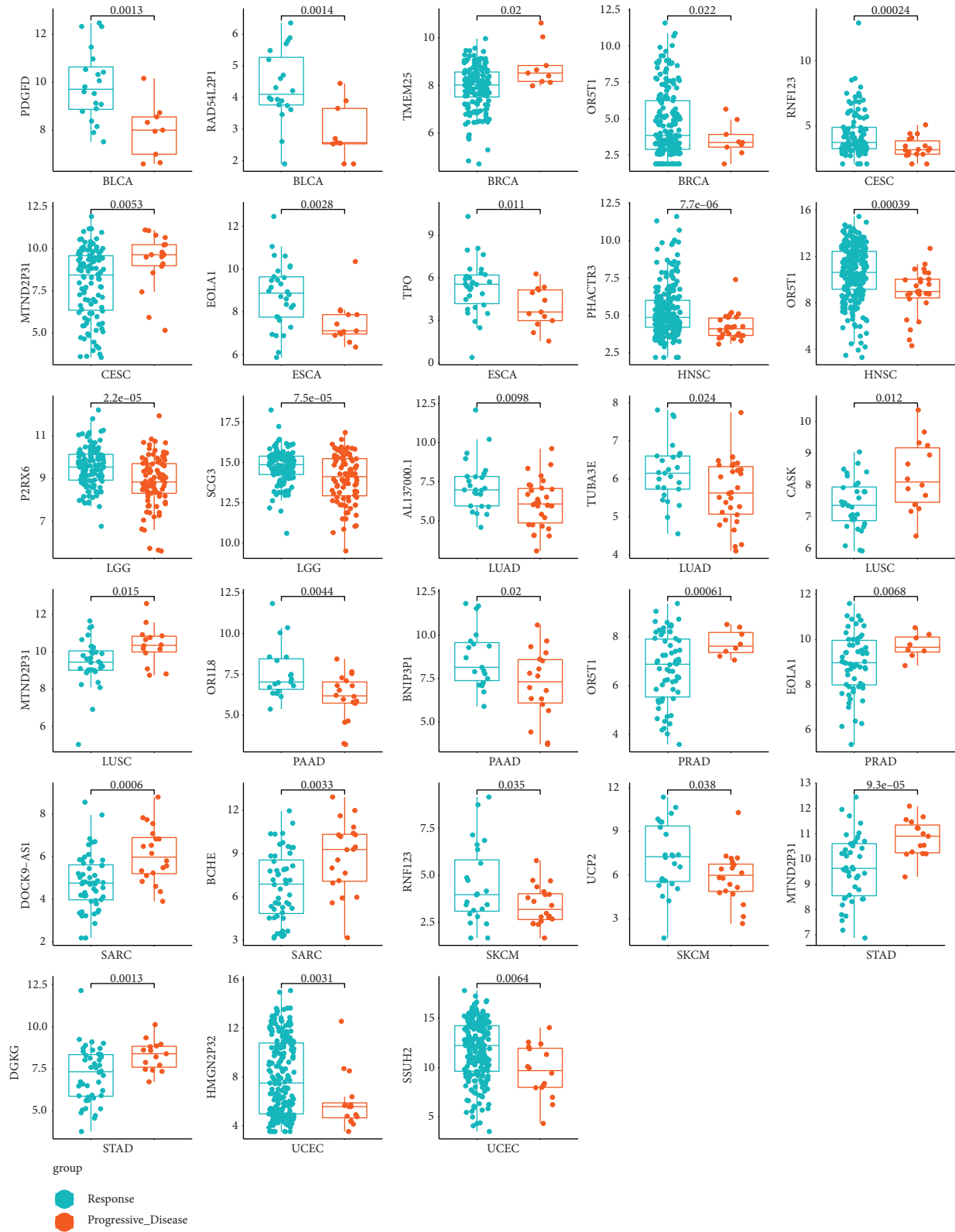


FIGURE 9: Boxplot of differentially expressed genes in different tumor types.

RNF123 ($p < 0.0001$), DTX1 ($p < 0.0001$), and CACNG2 ($p < 0.0001$) tended to be highly expressed in the *R* sample, and the survival rate of the high expression group was

significantly higher than that of the low expression group. These results revealed that *R* group patients had a better survival rate than PD patients.

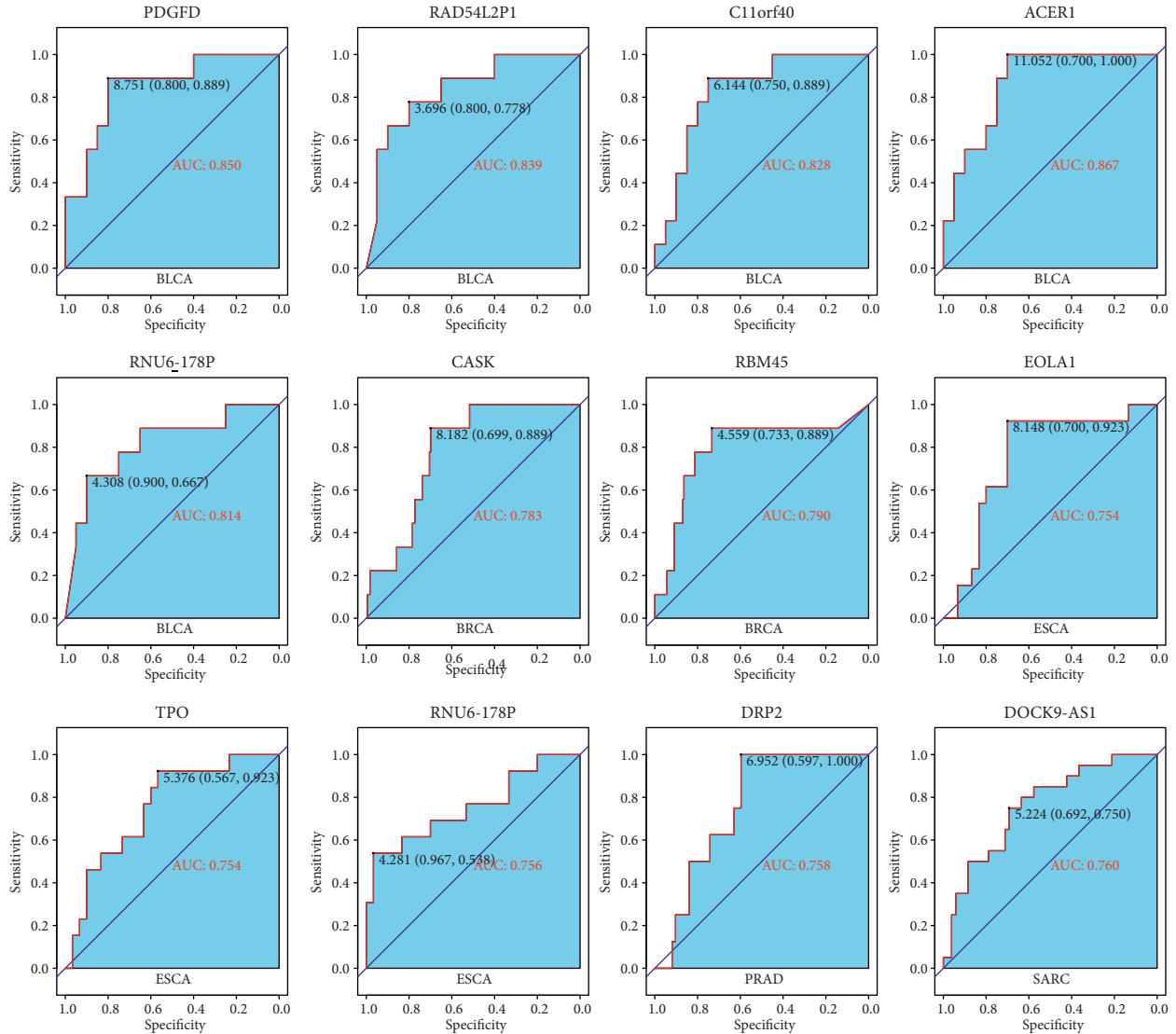


FIGURE 10: ROC map of differentially expressed genes in different tumor types.

3.5. Using Logistic Regression Model to Distinguish R and PD Samples. As mentioned above, the discriminative power of single-gene signature is not very strong. Therefore, we developed a binary logistic regression model to classify the samples. The prediction index was used to quantify each sample. We used 90% of the data set as the training set and 10% as the validation set. The ROC curve was drawn according to the prediction index of the regression model. Results are shown in Figure 7. The AUC value reached 0.822 and 0.787 in the training set and validation set at the best threshold score of 0.758, respectively (Figure 7(a)). Compared with the best AUC value of single-gene signature (OR1L8, AUC=0.700), the logistic regression model has better performance for distinguishing R and PD samples. The performance of classifying patients by prediction index is shown by visualizing predicted and actual classifications (Figure 7(b)). The result showed that PD samples were enriched on the left side of the prediction index curve and R samples are enriched on the right side of the prediction

index curve, implying that most of the samples can be classified correctly. We also assessed whether the prediction index can reach a better prediction for the survival rate of patients. As shown in Figure 7(c), the patients with a high prediction index had better survival than those patients with a low prediction index ($p < 0.0001$). This also implied that the RT responding patients had a better survival rate.

3.6. Performance of Single-Gene Signature in Different Tumor Types. In this study, the gene signature was obtained from 1664 TCGA clinical tumor samples. The tumors derive from 15 cancer types. However, given that there may be some degree of variation in expression patterns between different tumor types, it is necessary to select the optimal subset for each tumor type from the 100-gene signature. Therefore, we analyzed the discrimination ability of each gene for R and PD samples in each cancer type. Genes with a p value of less than 0.05 (p value less than 0.1 in few of tumor types) were

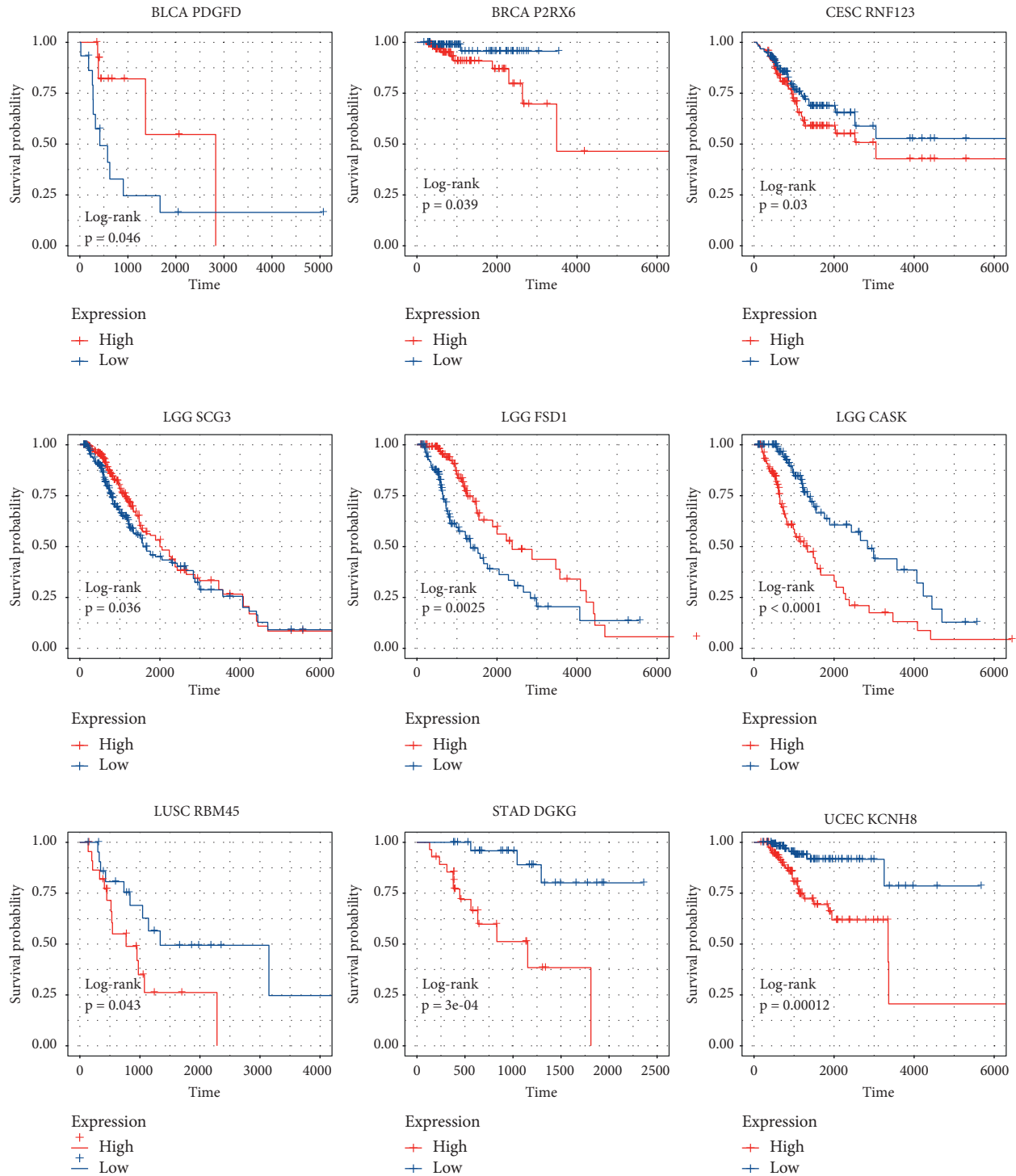


FIGURE 11: Overall survival curve of differentially expressed genes in different tumor types.

chosen to build tumor-specific gene subsets (Table S2). Cluster analysis revealed two distinct expression patterns of gene subsets in several cancer types, such as BLCA, PAAD, LUSC, and LGG (Figure 8). Many genes exhibited good ability to distinguish between *R* and *PD* groups (Figure 9). For example, RNF123 ($p = 0.00024$, in CESC), PHACTR3 ($p = 7.7 E - 06$, in HNSC), OR5T1 ($p = 0.00039$, in HNSC), P2RX6 ($p = 2.2 E - 05$, in LGG), and SCG3 ($p = 7.5 E - 05$, in

LGG) were expressed higher in *R* samples, while MTND2P31 ($p = 0.0053$, in CESC), OR5T1 ($p = 0.00061$, in PRAD), EOLA1 ($p = 0.0068$, in PRAD), BCHE ($p = 0.0033$, in SARC), MTND2P31 ($p = 9.3 E - 05$, in STAD), and DGKG ($p = 0.0013$, in STAD) were expressed higher in *PD* samples. Analysis by ROC curve also revealed several genes had high specificity of distinction (Figure 10), such as ACER1 (AUC = 0.867, in BLCA), RBM45 (AUC = 0.790, in BRCA),

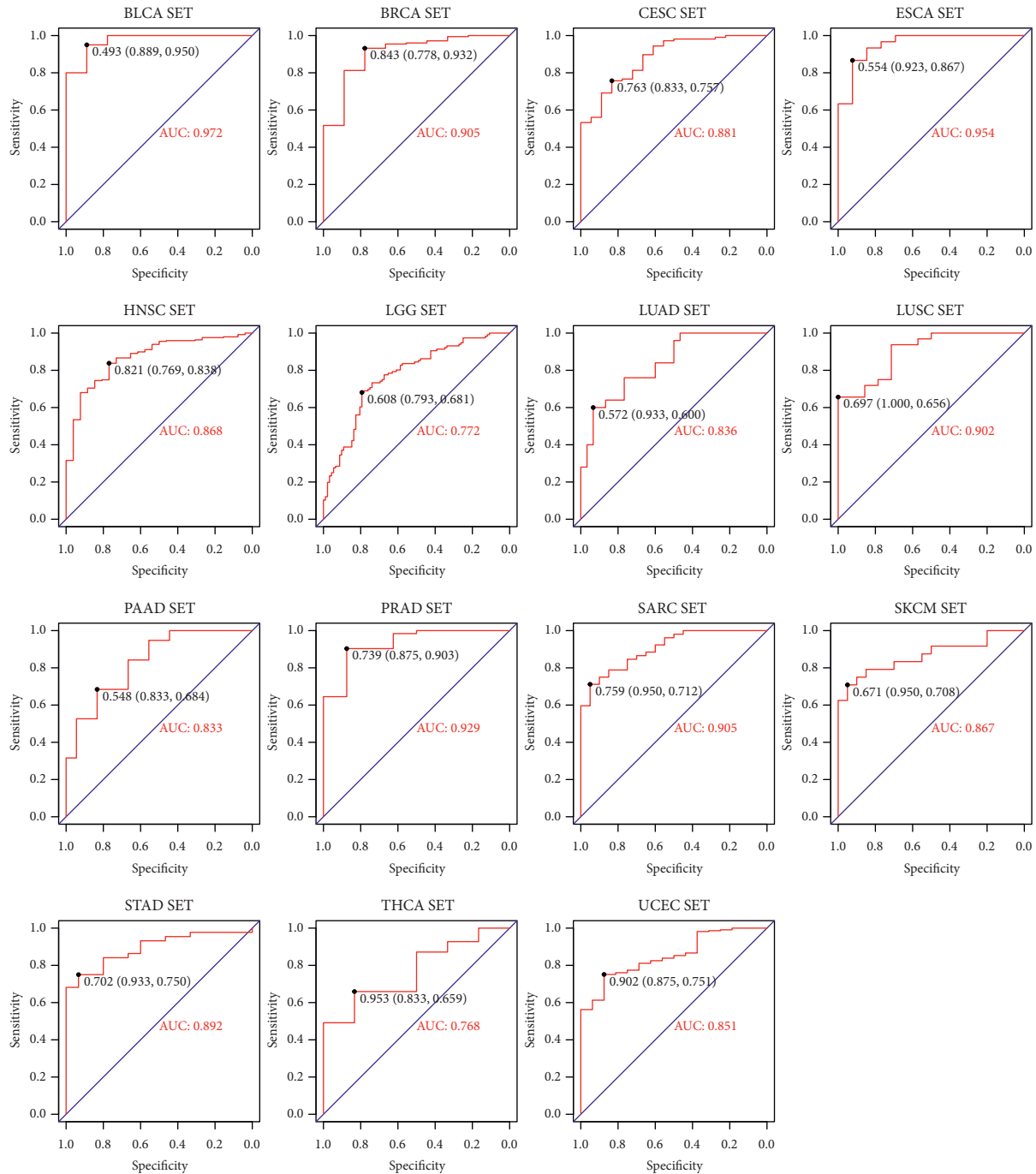


FIGURE 12: Logistic regression prediction ROC diagram of different tumor types.

RNU6 – 178P (AUC = 0.756, in ESCA), and DOCK9 – AS1 (AUC = 0.760, in SARC). Several genes also exhibited the predictive power for survival of patients (Figure 11). The higher expression of P2RX6 ($p = 0.039$, in BRCA), RNF123 ($p = 0.03$, in CESC), CASK ($p < 0.0001$, in LGG), DGKG ($p = 3E-04$, in STAD), RBM45 ($p = 0.043$, in LUSC), and KCN8 ($p = 0.00012$, in UCEC) was associated with worse survival, while the higher expression of PDGFD ($p = 0.046$, in BLCA), SCG3 ($p = 0.036$, in LGG) and FSD1 ($p = 0.0025$, in LGG) was associated with better survival.

3.7. Using Logistic Regression Model to Distinguish R and PD Samples in Different Tumor Types. Analysis of ROC curve revealed that the higher specificity was achieved in each cancer type by using the prediction index (Figure 12). For example, the AUC value of BLCA obtained by the prediction index was higher than from ACER1 (0.972 vs 0.867). Our results showed that most of the samples can be correctly classified in each cancer type (Figure 13). We also investigated whether the predictive model could provide better performance for the prediction of survival rate for each

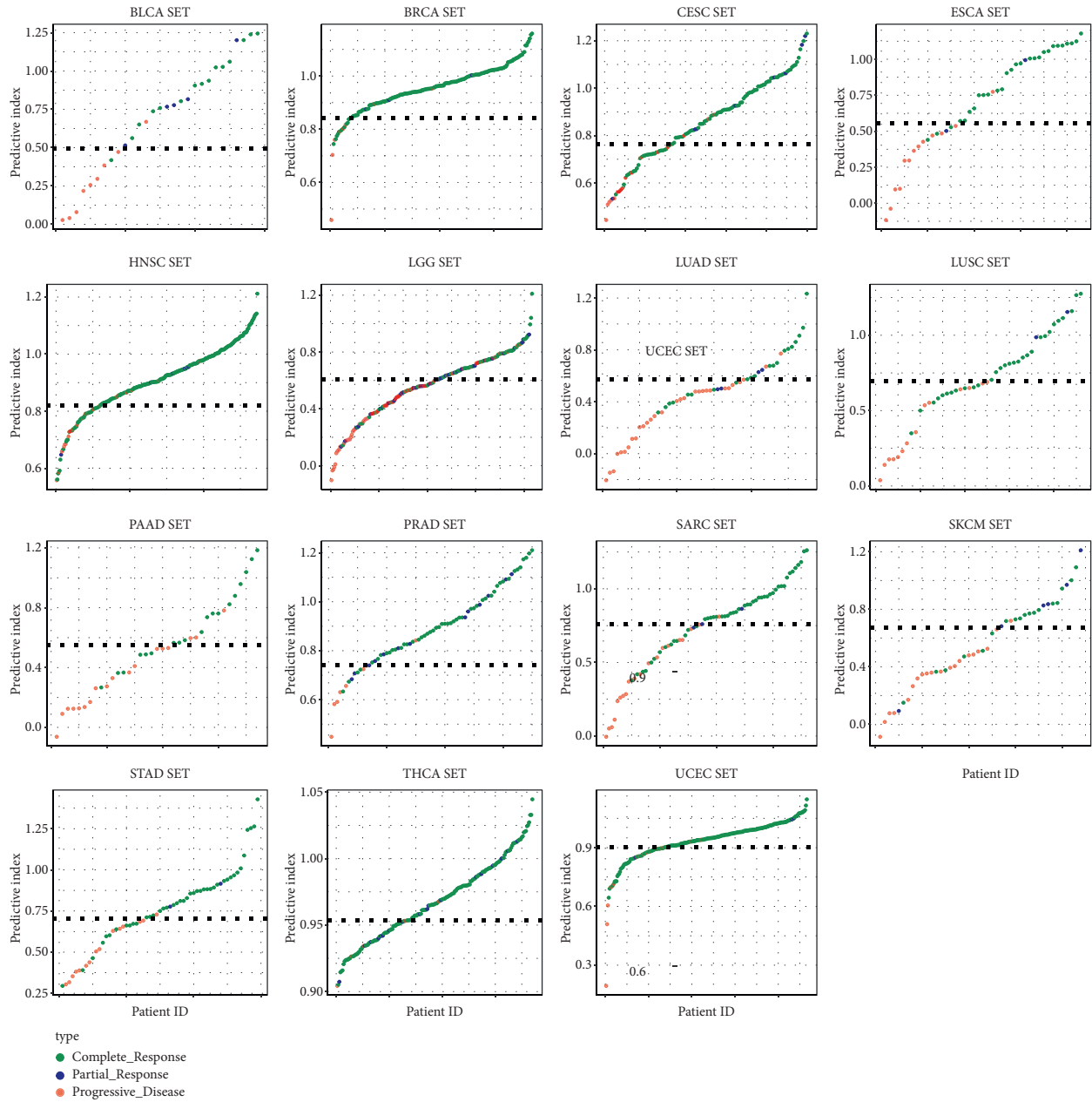


FIGURE 13: Logistic regression predictors of different tumor types.

cancer type. As shown in Figure 14, patients with a high prediction index were associated with better survival.

4. Discussion

In this study, we identified the differentially expressed genes between RT responder and nonresponder. Among them, MDC1 [26], UCP2 [27], and RBM45 [28] have been demonstrated to be involved in the DNA damage pathway and radiosensitivity. MDC1 [26] was identified as a component of DNA repair complex, controlling the damage-induced cell cycle arrest checkpoint. Cells lacking MDC1 are sensitive to ionizing radiation. Our study showed that MDC1 was expressed higher in the PD group, which might lead to

stronger reparability in PD patients. UCP2 [27] is a mitochondrial transporter, which can produce proton leakage on the inner membrane of mitochondria, thus uncoupling oxidative phosphorylation and ATP synthesis. The previous studies demonstrated that irradiation treatment can increase the expression level of UCP2, and silencing of UCP2 increased the radiosensitivity of HeLa cells and led to increased apoptosis, cell cycle arrest in G2/M, and mitochondrial ROS. Our results showed that the expression of UCP2 was significantly upregulated in the PD group, supporting that UCP2 has a role in radioresistance. RBM45 [28], also named DRB1, was recently found to be a FUS-interacting RBP. The previous study demonstrated that silencing of RBM45 led to a decreased efficiency in DSBs repair. Consistent with the

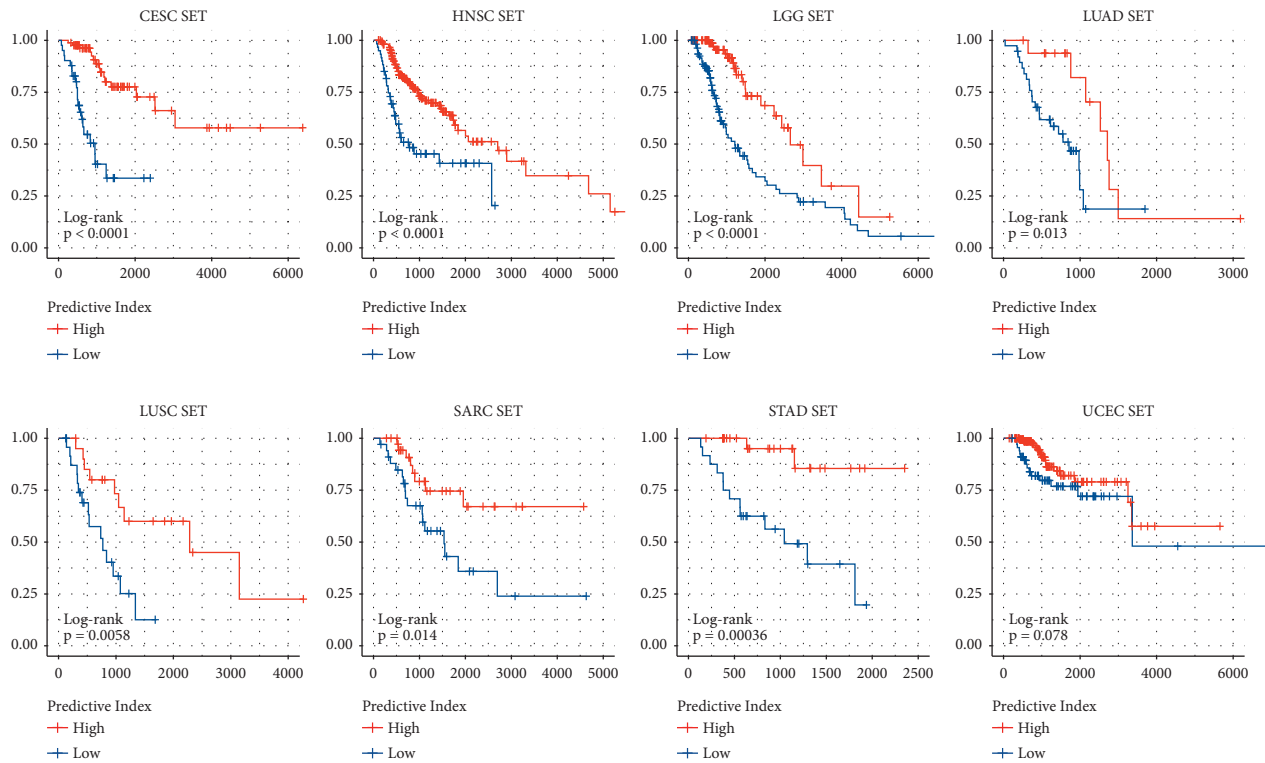


FIGURE 14: Survival curves of different tumor types.

results of the previous study, our results showed that RBM45 was expressed higher in the PD group.

Besides these three genes, several genes involving migration, cell proliferation, cell invasion, and EMT were identified in our study. For example, BCL9L [33], P2RX6 [34], RER1 [35], EFNA2 [36], CASK [37], CERCAM [38], and PTPRN [39] were demonstrated to promote EMT, metastasis, and invasion of tumor cells. Six genes involving ubiquitination proteasome process, including RBM45, TRIM9, PTPRN, RNF123, RNF220, and DTX1, were also identified. Ubiquitination plays an important role in innate and postnatal regulation of cell differentiation and cell survival. Previous studies have reported that ubiquitination-related proteins were involved in the regulation of DNA damage pathways and influenced the radiosensitivity of tumor cells [40, 41]. For example, Santra et al. found that E3 ubiquitin ligase FBXO31 mediates the degradation of cyclin D1 through ubiquitination and proteasome-mediated protein degradation pathway [40]. Knockdown of FBXO31 could prevent cells from undergoing efficient G1 arrest following γ -irradiation and greatly increased the sensitivity to DNA damage [41]. The differentially expressed genes also include several noncoding RNAs in our study. UFC1 is a long-chain noncoding RNA and was expressed higher in the PD group. Previous study showed that UFC1 was elevated and predicted poor prognosis of gastric cancer (GC), and knockdown of UFC1 inhibited the proliferation, migration, and invasion of GC cells [42].

Our study identified a series of genes that were differentially expressed in RT responders and nonresponders, providing useful clues for studying the molecular

mechanisms of tumor radioresistance. Those genes that were overexpressed in RT nonresponders could be potential targets for radiosensitization. The prediction models and gene signatures identified here also have the potential clinical application. A targeted mRNA sequencing technology based on the gene signatures identified in this study could be developed and used to detect mRNA levels in clinical samples. By analyzing the gene signatures, patients who will benefit from RT are identified, which could reduce the number of patients receiving unnecessary treatment and greatly reduce the cost of oncology treatment.

Abbreviations

RT:	Radiotherapy;
EMT:	Epithelial-mesenchymal transition
AUC:	Area under the curve;
DSBs:	DNA double-strand breaks
RNA-Seq:	High-throughput sequencing methods for mRNA;
TCGA:	The Cancer Genome Atlas
BLCA:	Bladder urothelial carcinoma;
BRCA:	Invasive breast cancer
CESC:	Cervical squamous cell carcinoma;
ESCA:	Esophageal cancer
HNSC:	Head and neck squamous cell carcinoma;
LGG:	Low-grade glioma
LUAD:	Lung adenocarcinoma;
LUSC:	Lung squamous cell carcinoma
PAAD:	Pancreatic cancer;
PRAD:	Prostate adenocarcinoma

SARC: Sarcoma;
 SKCM: Cutaneous melanoma
 STAD: Gastric adenocarcinoma;
 THCA: Thyroid carcinoma
 UCEC: Endometrial carcinoma;
 R: Response
 PD: Progressive disease;
 DEGs: Differentially expressed genes
 ROC: Receiver operating characteristic curve;
 DDR: DNA-damage response.

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.

Authors' Contributions

Yu Xu and Chao Tang contributed equally to this work.

Acknowledgments

This work was supported in part by projects (cstc2020jxjl130015, cstc2021jcyj-msxmX0767, cstc2019jcyj-msxmX0671, and cstc2020jcyj-msxmX1093) from the National Natural Science Foundation of China (no. 82073347) and the Integrated Innovation and Application of Key Technologies for Precise Prevention and Treatment of Primary Lung Cancer (no. 2019ZX002) from the Chongqing Municipal Health Committee.

Supplementary Materials

Table S1: Molecular functional annotation of top 100 differential genes in all cancer species analyzed. Table S2: DEGs used by cancer species to distinguish between R and PD samples. (*Supplementary Materials*)

References

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.
- [2] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [3] E. J. Moding, M. B. Kastan, and D. G. Kirsch, "Strategies for optimizing the response of cancer and normal tissues to radiation," *Nature Reviews Drug Discovery*, vol. 12, no. 7, pp. 526–542, 2013.
- [4] U. S. Srinivas, B. W. Q. Tan, B. A. Vellayappan, and A. D. Jeyasekharan, "ROS and the DNA damage response in cancer," *Redox Biology*, vol. 25, Article ID 101084, 2019.
- [5] L. Chang, P. Graham, J. Hao et al., "Cancer stem cells and signaling pathways in radioresistance," *Oncotarget*, vol. 7, no. 10, pp. 11002–11017, 2016.
- [6] K. Savitsky, A. Bar-Shira, S. Gilad et al., "A single ataxia telangiectasia gene with a product similar to PI-3 kinase," *Science*, vol. 268, no. 5218, pp. 1749–1753, 1995.
- [7] M. Ahmed, L. Li, C. Pinnix et al., "ATM mutation and radiosensitivity: an opportunity in the therapy of mantle cell lymphoma," *Critical Reviews In Oncology-Hematology*, vol. 107, pp. 14–19, 2016.
- [8] J. L. Bernstein, R. W. Haile, M. Stovall et al., "Radiation exposure, the ATM gene, and contralateral breast cancer in the women's environmental cancer and radiation epidemiology study," in *Journal of the National Cancer Institute: Journal of the National Cancer Institute*, P. Concannon, Ed., vol. 102, no. 7, pp. 475–483, 2010.
- [9] M. Choi, T. Kipps, and R. Kurzrock, "ATM mutations in cancer: therapeutic implications," *Molecular Cancer Therapeutics*, vol. 15, no. 8, pp. 1781–1791, 2016.
- [10] M. J. Kerins and A. Ooi, "A catalogue of somatic NRF2 gain-of-function mutations in cancer," *Scientific Reports*, vol. 8, no. 1, Article ID 12846, 2018.
- [11] Y. Jeong, N. T. Hoang, A. Lovejoy et al., "Role of KEAP1/NRF2 and TP53 mutations in lung squamous cell carcinoma development and radiation resistance," *Cancer Discovery*, vol. 7, no. 1, pp. 86–101, 2017.
- [12] Z. Gao, E. H. Sarsour, A. L. Kalen, L. Li, M. G. Kumar, and P. C. Goswami, "Late ROS accumulation and radiosensitivity in SOD1-overexpressing human glioma cells," *Free Radical Biology and Medicine*, vol. 45, no. 11, pp. 1501–1509, 2008.
- [13] X. Huang, S. Taeb, S. Jahangiri et al., "miRNA-95 mediates radioresistance in tumors by targeting the sphingolipid phosphatase SGPP1," *Cancer Research*, vol. 73, no. 23, pp. 6972–6986, 2013.
- [14] N. Mercatelli, V. Coppola, D. Bonci et al., "The inhibition of the highly expressed miR-221 and miR-222 impairs the growth of prostate carcinoma xenografts in mice," *PLoS One*, vol. 3, no. 12, Article ID e4029, 2008.
- [15] K. Hatano, B. Kumar, Y. Zhang et al., "A functional screen identifies miRNAs that inhibit DNA repair and sensitize prostate cancer cells to ionizing radiation," *Nucleic Acids Research*, vol. 43, no. 8, pp. 4075–4086, 2015.
- [16] C. G. Xu, M. F. Yang, J. X. Fan, and W. Wang, "MiR-30a and miR-205 are downregulated in hypoxia and modulate radiosensitivity of prostate cancer cells by inhibiting autophagy via TP53INP1," *European Review for Medical and Pharmacological Sciences*, vol. 20, no. 8, pp. 1501–1508, 2016.
- [17] Z. Tao, S. Xu, H. Ruan et al., "MiR-195/-16 family enhances radiotherapy via T cell activation in the tumor microenvironment by blocking the PD-L1 immune checkpoint," *Cellular Physiology and Biochemistry*, vol. 48, no. 2, pp. 801–814, 2018.
- [18] A. Mao, Y. Liu, Y. Wang et al., "miR-449a enhances radiosensitivity through modulating pRb/E2F1 in prostate cancer cells," *Tumor Biology*, vol. 37, no. 4, pp. 4831–4840, 2016.
- [19] Z. Xu, Y. Zhang, J. Ding et al., "miR-17-3p downregulates mitochondrial antioxidant enzymes and enhances the radiosensitivity of prostate cancer cells," *Molecular Therapy - Nucleic Acids*, vol. 13, pp. 64–77, 2018.
- [20] A. C. Mueller, D. Sun, and A. Dutta, "The miR-99 family regulates the DNA damage response through its target SNF2H," *Oncogene*, vol. 32, no. 9, pp. 1164–1172, 2013.

- [21] X. Zhu, Y. Wang, L. Tan, and X. Fu, "The pivotal role of DNA methylation in the radio-sensitivity of tumor radiotherapy," *Cancer Medicine*, vol. 7, no. 8, pp. 3812–3819, 2018.
- [22] Z. Liu, H. Chen, Y. Xia, A. L. Kwan, and Z. Chen, "Relationship between methylation status of ERCC1 promoter and radio-sensitivity in glioma cell lines," *Cell Biology International*, vol. 5, no. 10, pp. 156–159, 2007.
- [23] Z. Wang, M. Gerstein, and M. Snyder, "RNA-Seq: a revolutionary tool for transcriptomics," *Nature Reviews Genetics*, vol. 10, no. 1, pp. 57–63, 2009.
- [24] J. Costa-Silva, D. Domingues, and F. M. Lopes, "RNA-Seq differential expression analysis: an extended review and a software tool," *PLoS One*, vol. 12, no. 12, Article ID e0190152, 2017.
- [25] Q. Q. Wang, S. C. Yu, X. Qi et al., "Overview of logistic regression model analysis and application," *Zhonghua yu fang yi xue za zhi [Chinese journal of preventive medicine]*, vol. 53, no. 9, pp. 955–960, 2019.
- [26] G. S. Stewart, B. Wang, C. R. Bignell, A. M. R. Taylor, and S. J. Elledge, "MDC1 is a mediator of the mammalian DNA damage checkpoint," *Nature*, vol. 421, no. 6926, pp. 961–966, 2003.
- [27] C. H. Liu, Z. H. Huang, X. Y. Dong et al., "Inhibition of uncoupling protein 2 enhances the radiosensitivity of cervical cancer cells by promoting the production of reactive oxygen species," *Oxidative Medicine and Cellular Longevity*, vol. 2020, pp. 1–13, Article ID 5135893, 2020.
- [28] J. Gong, M. Huang, F. Wang et al., "RBM45 competes with HDAC1 for binding to FUS in response to DNA damage," *Nucleic Acids Research*, vol. 45, no. 22, pp. 12862–12876, 2017.
- [29] S. Venuto and G. Merla, "E3 ubiquitin ligase TRIM proteins, cell cycle and mitosis," *Cells*, vol. 8, no. 5, pp. 510–525, 2019.
- [30] Y. Wang, M. Cui, X. Cai et al., "The oncoprotein HBXIP up-regulates SCG3 through modulating E2F1 and miR-509-3p in hepatoma cells," *Cancer Letters*, vol. 352, no. 2, pp. 169–178, 2014.
- [31] L. Di, M. Gu, Y. Wu et al., "SNAP25 is a potential prognostic biomarker for prostate cancer," *Cancer Cell International*, vol. 22, no. 1, 144 pages, 2022.
- [32] S. Nakasone, A. Suzuki, H. Okazaki et al., "Predictive markers based on transcriptome modules for vinorelbine-based adjuvant chemotherapy for lung adenocarcinoma patients," *Lung Cancer*, vol. 158, pp. 115–125, 2021.
- [33] X. Wang, M. Feng, T. Xiao et al., "BCL9/BCL9L promotes tumorigenicity through immune-dependent and independent mechanisms in triple negative breast cancer," *Oncogene*, vol. 40, no. 16, pp. 2982–2997, 2021.
- [34] D. Gong, J. Zhang, Y. Chen et al., "The m6A-suppressed P2RX6 activation promotes renal cancer cells migration and invasion through ATP-induced Ca²⁺ influx modulating ERK1/2 phosphorylation and MMP9 signaling pathway," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, p. 233, 2019.
- [35] S. Chen, J. Zhang, J. Chen et al., "RER1 enhances carcinogenesis and stemness of pancreatic cancer under hypoxic environment," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, 15 pages, 2019.
- [36] Y. Zhao, C. Cai, M. Zhang et al., "Ephrin-A2 promotes prostate cancer metastasis by enhancing angiogenesis and promoting EMT," *Journal of Cancer Research and Clinical Oncology*, vol. 147, no. 7, pp. 2013–2023, 2021.
- [37] J. Qu, Y. Zhou, Y. Li, J. Yu, and W. Wang, "CASK regulates Notch pathway and functions as a tumor promoter in pancreatic cancer," *Archives of Biochemistry and Biophysics*, vol. 701, Article ID 108789, 2021.
- [38] Y. Zuo, X. Xu, M. Chen, and L. Qi, "The oncogenic role of the cerebral endothelial cell adhesion molecule (CERCAM) in bladder cancer cells in vitro and in vivo," *Cancer Medicine*, vol. 10, no. 13, pp. 4437–4450, 2021.
- [39] X. Song, X. Jiao, H. Yan et al., "Overexpression of PTPRN promotes metastasis of lung adenocarcinoma and suppresses NK cell cytotoxicity," *Frontiers in Cell and Developmental Biology*, vol. 9, Article ID 622018, 2021.
- [40] M. K. Santra, N. Wajapeyee, and M. R. Green, "F-box protein FBXO31 mediates cyclin D1 degradation to induce G1 arrest after DNA damage," *Nature*, vol. 459, no. 7247, pp. 722–725, 2009.
- [41] Y. Lu, J. Li, D. Cheng et al., "The F-box protein FBXO44 mediates BRCA1 ubiquitination and degradation," *Journal of Biological Chemistry*, vol. 287, no. 49, pp. 41014–41022, 2012.
- [42] X. Zhang, W. Liang, J. Liu et al., "Long non-coding RNA UFC1 promotes gastric cancer progression by regulating miR-498/Lin28b," *Journal of Experimental & Clinical Cancer Research*, vol. 37, no. 1, p. 134, 2018.

Research Article

Detection Anomaly in Video Based on Deep Support Vector Data Description

Bokun Wang,¹ Caiqian Yang^{1b},² and Yaojing Chen³

¹College of Civil Engineering and Mechanics, Xiangtan University, Xiangtan 411100, China

²School of Civil Engineering, Southeast University, Nanjing 210096, China

³Jiangsu Expressway Engineering Maintenance Co., Ltd., Huaian 223005, China

Correspondence should be addressed to Caiqian Yang; ycqjxx@hotmail.com

Received 3 April 2022; Accepted 18 April 2022; Published 4 May 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Bokun 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.

Video surveillance systems have been widely deployed in public places such as shopping malls, hospitals, banks, and streets to improve the safety of public life and assets. In most cases, how to detect video abnormal events in a timely and accurate manner is the main goal of social public safety risk prevention and control. Due to the ambiguity of anomaly definition, the scarcity of anomalous data, as well as the complex environmental background and human behavior, video anomaly detection is a major problem in the field of computer vision. Existing anomaly detection methods based on deep learning often use trained networks to extract features. These methods are based on existing network structures, instead of designing networks for the goal of anomaly detection. This paper proposed a method based on Deep Support Vector Data Description (DSVDD). By learning a deep neural network, the input normal sample space can be mapped to the smallest hypersphere. Through DSVDD, not only can the smallest size data hypersphere be found to establish SVDD but also useful data feature representations and normal models can be learned. In the test, the samples mapped inside the hypersphere are judged as normal, while the samples mapped outside the hypersphere are judged as abnormal. The proposed method achieves 86.84% and 73.2% frame-level AUC on the CUHK Avenue and ShanghaiTech Campus datasets, respectively. By comparison, the detection results achieved by the proposed method are better than those achieved by the existing state-of-the-art methods.

1. Introduction

In order to improve the safety of public life and assets, video surveillance systems have been widely deployed in public places such as shopping malls, hospitals, banks, and streets. In most cases, how to detect video abnormal events in a timely and accurate manner is the main goal of social public safety risk prevention and control. Video abnormal events are defined as abnormal or irregular patterns in the video that do not conform to normal patterns. These incidents often include fights, riots, violations of traffic rules, trampling, holding arms, and abandoning luggage. However, due to the ambiguity of anomaly definitions, the scarcity of anomalous data, and the complex environmental background and human behavior, video anomaly detection is a major problem in the field of computer vision. In a nutshell,

most of the current research work on video anomaly detection can be divided into two steps, such as feature extraction and normal model training [1]. Feature extraction can be achieved by manual technology or automatic feature extraction technology (representation learning or features based on deep learning). In normal model training, normal samples are used for learning, and then samples that do not conform to the learned model are judged as abnormal events. Then, the classification according to features can be divided into three different methods [2]. The first type is the trajectory-based methods [3]. This type of method obtains trajectory features by tracking the target. However, in dense scenes, the target tracking is a big problem. The second type of methods is based on global features [4, 5]. This type of method takes the video frame as a whole and extracts some low-level or middle-level features such as spatiotemporal

gradients and optical flow. In a moderately crowded and dense environment, these methods can keep effective. The third type is the grid feature-based methods [6]. This type of method often divides the video frame into multiple small grids through dense sampling and then extracts the underlying features of a single grid because each grid can be individually evaluated. According to different normal model training methods, the present methods can also be divided into three different types. The first type is the cluster-based method [7]. This type of method is often based on an assumption that the normal sample belongs to a category or is relatively far from the cluster center. The abnormal samples do not belong to any category or are far away from the cluster center and then cluster the normal samples to build the model. The second is the method based on sparse reconstruction [8, 9]. This type of method assumes that the sparse linear combination of patterns can represent normal activities with the smallest reconstruction error. Because there is no abnormal activity in the training data set, it can represent abnormal patterns with a large reconstruction error. The third type is probabilistic model-based methods. This method considers that normal samples conform to a certain probability distribution, while abnormal samples do not conform to this distribution.

Recently, the latest progress of deep learning has proved the obvious advantages of deep learning-based methods in many computer vision applications [10]. As one of the tasks in computer vision, video anomaly detection is no exception. Different from traditional manual feature-based methods, deep learning methods often use pretrained networks to extract high-level features from videos or use existing network structures to establish end-to-end anomaly detection models based on normal models. For the former idea [11, 12], there is not much difference between the two steps of traditional abnormal event detection. For the latter idea [13–16], the two steps of feature extraction and model building are often jointly optimized in a deep network. These end-to-end deep networks include Auto-Encoder (AE), Deep Siamese Network (DSN), and Generative Adversarial Nets (GANs) [14, 17–20]. However, these network models are often designed for other tasks such as generative models and compression, rather than for anomaly detection tasks.

In the framework of deep learning, this paper proposes a new anomaly detection method based on Deep Support Vector Data Description (DSVDD) for anomaly detection tasks. Through DSVDD, not only can the smallest size data hypersphere be found to establish SVDD but also useful data feature representations and normal models can be learned. To this end, DSVDD uses a jointly trained deep neural network to map normal sample data to the smallest volume hypersphere. Then, in the test, the samples mapped inside the hypersphere are judged as normal, while the samples mapped outside the hypersphere are judged as abnormal. The RGB graph and the optical flow graph are composed of a 6-channel data and directly input into a DSVDD model; that is, it can detect the appearance abnormality and movement abnormality at the same time. The experimental results on the two public data sets of Avenue [9] and ShanghaiTec [17]

show that the detection results of the method proposed in this paper are excellent, which exceed the state of the art.

2. Principle of Algorithm

The overall process of the method proposed in this paper is shown in Figure 1. In the training phase, the RGB images and optical flow diagrams of the training samples are intensively sampled and then merged into a 6-channel data to train the DSVDD model. In the testing phase, the RGB image and optical flow diagram composition of the video frame to be tested are also obtained after inputting the 6-channel data into the learned DSVDD model. It is determined whether the area is abnormal. In this section, the principle of SVDD is first briefly introduced, and then the training and testing process of video abnormal events based on DSVDD is described.

2.1. SVDD. SVDD is a description method based on boundary data (support vector). Its goal is to find a hypersphere that contains all or almost all training samples and has the smallest volume (the center is $c \in \mathcal{F}_k$, and the radius is $R > 0$). In fact, the SVDD optimization problem can be transformed as follows:

$$\begin{aligned} \min_{R, \mathcal{W}} R^2 + \frac{1}{\nu n} \sum_i \xi_i \\ \text{s.t. } \|\phi_k(x_i) - c\|_{\mathcal{F}_k}^2 \leq R^2 + \xi_i, \quad \xi_i \geq 0. \end{aligned} \quad (1)$$

In (1), the slack variable allows a soft boundary; $\xi_i \geq 0$ and $\nu \in (0, 1]$ are hyperparameters to control the balance between the penalty term and the volume edge of the hypersphere. Therefore, a point that falls outside the hypersphere, such as $\|\phi_k(x_i) - c\|_{\mathcal{F}_k}^2 > R^2$, is decided to be abnormal. SVDD has been widely used in fields such as anomaly detection, face recognition, speech recognition, image restoration, and medical imaging [21].

2.2. DSVDD. DSVDD learns a deep neural network $\phi(\cdot; \mathcal{W})$ with the weight \mathcal{W} , so that the input normal sample space can be mapped to a hypersphere with the center and radius of the smallest. The normal sample is mapped in the hypersphere $\mathcal{X} \subseteq \mathbb{R}^d$, and the abnormal sample is mapped on the hypersphere.

Specifically, for the sample area input space $\mathcal{X} \subseteq \mathbb{R}^d$ and output space $\mathcal{F} \subseteq \mathbb{R}^p$, a neural network with $L \in \mathbb{N}$ hidden layers can project the input space to the output space $\mathcal{X} \rightarrow \mathcal{F}$, where $\mathcal{W} = \{W^1, W^2, \dots, W^L\}$ are the weights of the hidden layers $\ell = \{1, 2, \dots, L\}$ correspondingly. Therefore, $\phi(x; \mathcal{W}) \in \mathcal{F}$ is the characteristic representation of the input sample $x \in \mathcal{X}$. The goal of the DSVDD method is to jointly optimize the network weights \mathcal{W} and the output space to meet the minimum hyperspherical constraints of the center c and the radius R . Then, given the training sample $\mathcal{D}_n = \{x_1, x_2, \dots, x_n\}$, the soft-boundary objective function of DSVDD is as follows:

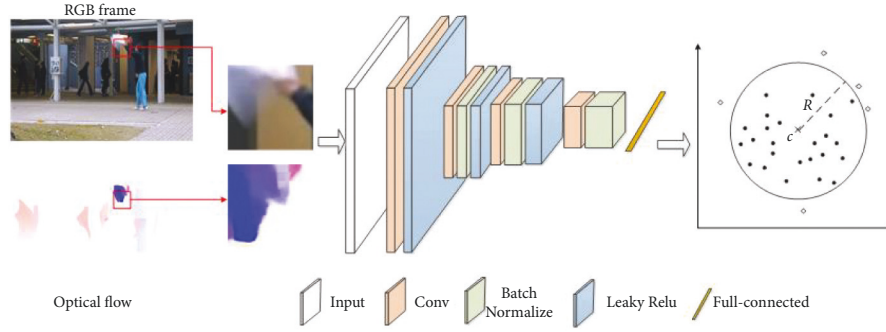


FIGURE 1: The flow chart of video anomaly detection based on DSVDD.

$$\min_{R, \mathcal{W}} R^2 + \frac{1}{vn} \sum_{i=1}^n \max \left\{ 0, \|\phi(x_i; \mathcal{W}) - c\|^2 - R^2 \right\} + \frac{\lambda}{2} \sum_{\ell=1}^L \|\mathcal{W}\|_F^2. \quad (2)$$

For (2), in the SVDD method, the minimization of R^2 means to minimize the volume of the hypersphere. The second item is the penalty items that are mapped out of the hypersphere through the neural network, such as those from the center of the hypersphere $\|\phi(x_i; \mathcal{W}) - c\|$ greater than radius R . The hyperparameter $v \in (0, 1]$ controls the balance between the volume of the hypersphere and the deviation of the boundary, which allows certain points to be mapped to the outside of the sphere. The last item is the network parameter weight \mathcal{W} the attenuation regularization term, where $\lambda > 0$ and $\|\cdot\|_F$ represents the Frobenius norm.

The optimization of (2) enables the network to learn weights \mathcal{W} , so that the data points can be closely projected to the center of the hypersphere c nearby. For this reason, the deep network must extract the common factors of data changes. In fact, normal samples can often be mapped closer to the center of the hypersphere c , while abnormal samples are mapped farther from the center or outside the hypersphere. In this way, a compact description of the normal model is obtained.

In actual tasks, it is often assumed that the training samples are all normal samples, so the objective function can be simplified to a single-class classification problem as follows:

$$\min_{\mathcal{W}} \frac{1}{n} \sum_{i=1}^n \|\phi(x_i; \mathcal{W}) - c\|^2 + \frac{\lambda}{2} \sum_{\ell=1}^L \|\mathcal{W}\|_F^2. \quad (3)$$

DSVDD simply uses a secondary loss to punish the distance of each deep network representation $\phi(x_i; \mathcal{W})$ and c . The second term is the regularization term of network parameter weight attenuation \mathcal{W} , $\lambda > 0$. Equation (3) can also be regarded as a hypersphere with the smallest volume as the center. However, unlike Equation (2) using a soft boundary, Equation (3) shrinks the sphere by minimizing the average distance from the center of all data representations, instead of directly penalizing the radius and data representation that falls outside the sphere. Similarly, in order to map the samples as close to the center of the

hypersphere as possible, the deep neural network must extract the changing common factors.

The weights \mathcal{W} of the neural network in DSVDD can be optimized by common back propagation methods (such as stochastic gradient descent). Because the network weight \mathcal{W} and hypersphere radius R are with different scales, it is impossible to optimize DSVDD with one learning rate. Therefore, it is necessary to alternately optimize the network weights \mathcal{W} and hypersphere radius R by the alternate minimization/block coordinate descent method.

2.3. Test Phase. Given test sample area $x' \in \mathcal{X}$, the anomaly score can e calculated as follows:

$$s(x') = \|\phi(x'; \mathcal{W}^*) - c\|^2. \quad (4)$$

where \mathcal{W}^* are the trained network model parameters. It is worth noting that network parameters can fully describe the DSVDD model. And predictions can be made without storing any data, so DSVDD has a very low storage complexity. Therefore, the computational complexity during testing is small.

In order to infer whether the test sample area is an abnormal sample, thresholds can be set on $s(x')$ to make judgments as follows:

$$s(x') \underset{\text{normal}}{\overset{\text{abnormal}}{\geq}} \theta, \quad (5)$$

where θ is the threshold that determines the sensitivity of the detection method in this paper.

3. Experiment

3.1. Dataset. This paper evaluates the performance of the DSVDD method on two publicly available data sets, i.e., the Avenue data set [9] and ShanghaiTech data set [17]. The Avenue data set is one of the most widely used benchmarks for video anomaly detection. It contains 16 training video clips and 21 test video clips, including 47 abnormal incidents that occurred on the streets of the Chinese University of Hong Kong. Each video is about 1 minute long and has a resolution of 640×360 . Normal events are walking on the street, and abnormal events include running, loitering, and throwing. ShanghaiTech data set [17] is one of the largest newly proposed datasets for video anomaly detection.

TABLE 1: AUC scores of the anomaly detection results.

Method	Avenue	ShanghaiTech campus
Conv-AE	80.0%	60.9%
Stacked RNN	81.7%	68.0%
Unmasking	80.6%	—
Davide et al.	—	72.8%
Object-centric auto-encoders	86.5%	78.5%
MemAE	83.3%	72.2%
New baseline	85.1%	72.8%
Ours	86.8%	73.2%

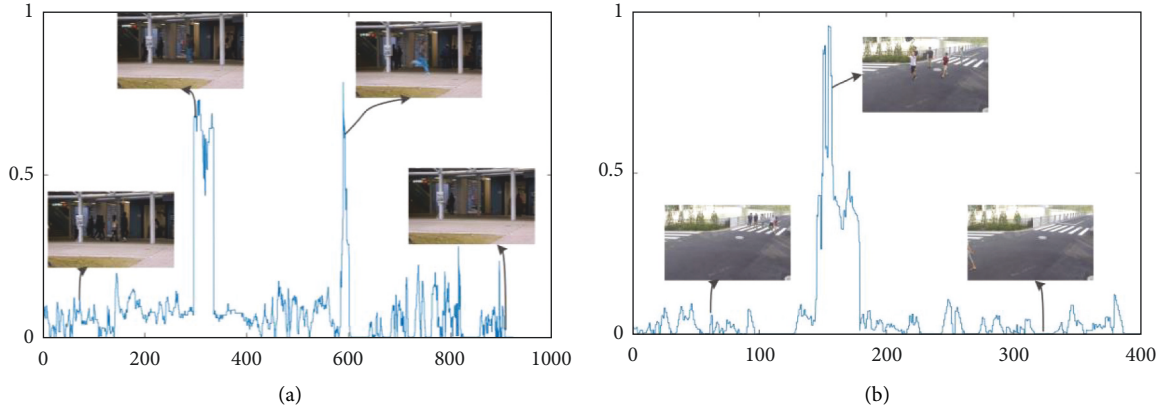


FIGURE 2: Illustration of detection results on different datasets. (a) Example of detection result on the Avenue data set (b) Example of detection result on the Avenue data set.

Unlike other data sets, the video clips in this data set come from 13 different cameras with different lighting conditions and camera angles. It has 330 training video clips and 107 test video clips containing 130 abnormal events. The resolution of the video frame is 856×480 . Abnormal events in this data set include chasing and noise.

3.2. Evaluation Index. According to previous work [14], this paper calculates the frame-level receiver operating characteristic (ROC) curve and uses the area under the curve (AUC) score as an evaluation indicator. A higher AUC score indicates better anomaly detection performance. If an area in the video frame is judged to be abnormal, the frame is judged to be abnormal. We first obtain the anomaly scores of all video frames and then calculate the frame-level AUC scores.

3.3. Supplementary Details. For the two data sets, each frame is adjusted to a size of 320×240 , and the optical flow image is calculated by the RAFT optical flow method provided in [22] through a network pretrained on the things data set. The original video frame and the calculated optical flow graph are combined into a 6-channel data, then cropped into 16×12 grid images according to the size of 20×20 , and then input into DSVDD for training and prediction. The deep neural network part of DSVDD is in accordance with Conv (16, 3×3)-Leaky ReLU-ConvTran ($32, 3 \times 3$)-BN-Leaky ReLU-ConvTran ($64, 3 \times 3$)-BN-Leaky ReLU-FullyConnectd64 structure. In the training phase, the batch size is set to 128,

the initial learning rate is 0.0003, the weight decay is 0.0001, and the training is performed 1000 iterations.

3.4. Experimental Results. This section compares the proposed DSVDD with the results obtained by several latest methods that only use positive sample training. These methods include Conv-AE [14], Stacked RNN [17], Unmasking [18], Davide et al. [19], Object-centric auto-encoder [20], MemAE [23], and New Baseline [24]. Table 1 lists the evaluation results of these methods on the frame-level anomaly detection on the two data sets [25–27].

On the Avenue dataset, the DSVDD method proposed in this paper is superior to the results obtained by other methods, with an AUC score of 87.4%, which is 2.3% higher than the baseline method proposed in 2018 [24]. As far as we know, in terms of the frame-level AUC scores of all test videos in this data set, the DSVDD proposed in this paper has achieved the best results. It is worth noting that the Object-centric auto-encoder [20] method achieved 89.3% of the frame-level AUC in their paper, but this is calculated through different indicators in their paper and the actual calculation of Object-centric. The frame-level AUC score obtained by the auto-encoder [20] method should be 86.5%, which is 0.9% lower than the method proposed in this paper.

On the ShanghaiTech dataset, the method DSVDD proposed in this paper achieves a frame-level AUC score of 74.5%, which is 1.7% higher than the baseline method proposed in 2018 [24] and second only to Object-centric auto-encoder [20]. The method achieved 78.5%. The Object-

centric auto-encoder [20] method uses an object detection-based method for anomaly detection, and its performance largely depends on the output of its object detection algorithm. Therefore, detection-based methods cannot determine abnormal events that have not occurred before, and this often occurs in abnormal detection. Similarly, the MemAE method [23] requires the help of a pretrained pose estimator to achieve better results, so it is limited to detecting abnormal events related to people. In contrast, the DSVDD method proposed in this article does not have this limitation and is very reliable when applied to various scenarios. Obviously, in addition to these two specially limited methods, the DSVDD method proposed in this paper is at least 1.7% ahead of other methods in frame-level AUC.

In Figure 2, some examples of abnormal score curves in the method proposed in this paper are shown, and some key frames with normal or abnormal events are given. Among them, the abscissa is the number of video frames, and the ordinate anomaly score has been normalized to 1. It can be seen that in the two data sets, the method proposed in this paper can correctly distinguish between normal and abnormal events. If an abnormal event occurs suddenly, such as running as shown in Figure 2(a), the abnormal score will increase sharply. If the abnormal event occurs slowly, as shown in Figure 2(b), the abnormal score will gradually increase. If the object that caused the abnormality disappears from the camera's field of view, the abnormality score will quickly decrease to close to 0.

4. Conclusion

In this paper, a video anomaly detection method based on DSVDD is proposed. DSVDD can be seen as a combination of deep learning and SVDD. It uses a jointly trained deep neural network to map normal sample data to the smallest volume hypersphere. Then, in the test, the samples mapped inside the hypersphere are judged as normal, while the samples mapped outside the hypersphere are judged as abnormal. A large number of experimental results on two public data sets show that the proposed method is significantly better than the existing methods, which proves the effectiveness of the anomaly detection method proposed in this paper. In the future, we will reduce the computational complexity on the basis of ensuring the accuracy of the algorithm and focus on improving the real-time performance of the algorithm to better apply it to actual scenarios. [25–27].

Data Availability

The datasets used in this paper can be accessed upon request.

Conflicts of Interest

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

References

- [1] W. Luo, W. Liu, D. Lian et al., "Video anomaly detection with sparse coding inspired deep neural networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1–15, 2019.

- [2] R. Nayak, U. C. Pati, and S. K. Das, "A comprehensive review on deep learning-based methods for video anomaly detection," *Image and Vision Computing*, vol. 106, Article ID 104078, 2021.
- [3] S. Cosar, G. Donatiello, V. Bogorny, C. Garate, L. O. Alvares, and F. Bremond, "Toward abnormal trajectory and event detection in video surveillance," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 3, pp. 683–695, 2017.
- [4] K. Xu, X. Jiang, and T. Sun, "Anomaly detection based on stacked sparse coding with i classification strategy," *IEEE Transactions on Multimedia*, vol. 20, no. 5, pp. 1062–1074, 2018.
- [5] K. W. Cheng, Y. T. Chen, and W. H. Fang, "Video anomaly detection and localization using hierarchical feature representation and Gaussian process regression[C]," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2909–2917, 2015.
- [6] R. Leyva, V. Sanchez, and C.-T. Li, "Video anomaly detection with compact feature sets for online performance," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3463–3478, 2017.
- [7] T. Zhao, F. Li, and P. Tian, "A deep-learning method for device activity detection in mMTC under imperfect CSI based on variational-autoencoder," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7981–7986, 2020.
- [8] D. Xu, Y. Yan, E. Ricci, and N. Sebe, "Detecting anomalous events in videos by learning deep representations of appearance and motion," *Computer Vision and Image Understanding*, vol. 156, pp. 117–127, 2017.
- [9] C. Lu, J. Shi, and J. Jia, "Abnormal event detection at 150 FPS in MATLAB," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2720–2727, Sydney, NSW, Australia, December 2013.
- [10] Y. Zerrouki, F. Harrou, N. Zerrouki, A. Dairi, and Y. Sun, "Desertification detection using an improved variational autoencoder-based approach through ETM-landsat satellite data," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 202–213, 2021.
- [11] X. Zhang, B. Hu, and Z. Pan, "Tensor representation based target detection for hyperspectral imagery," *Optical and Precision Engineering*, vol. 27, no. 2, pp. 488–498, 2019.
- [12] M. Ravanbakhsh, M. Nabi, H. Mousavi, E. Sanginetto, and N. Sebe, "Plug-and-Play CNN for crowd motion analysis:an application in abnormal event detection," in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, pp. 1689–1698, Lake Tahoe, NV, USA, March 2018.
- [13] Y. Wang, B. Dai, G. Hua, J. Aston, and D. Wipf, "Recurrent variational autoencoders for learning nonlinear generative models in the presence of outliers," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 6, pp. 1615–1627, 2018.
- [14] M. Hasan, J. Choi, J. Neumann, A. K. Roy-Chowdhury, and L. S. Davis, "Learning temporal regularity in video sequences," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, Las Vegas, NV, June 2016.
- [15] B. Ramachandra, M. J. Jones, and R. R. Vatsavai, "Learning a distance function with a Siamese network to localize anomalies in videos," in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, pp. 2598–2607, Snowmass Village, CO, March 2020.

- [16] X. Wang, Z. Che, B. Jiang et al., “Robust unsupervised video anomaly detection by multi-path frame prediction,” 2020, <https://arxiv.org/abs/2011.02763>.
- [17] W. Luo, W. Liu, and S. Gao, “A revisit of sparse coding based anomaly detection in stacked rnn framework,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 341–349, Venice, Italy, October 2017.
- [18] R. T. Ionescu, S. Smeureanu, B. Alexe, and M. Popescu, “Unmasking the abnormal events in video,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2895–2903, Venice, Italy, October 2017.
- [19] D. Abati, A. Porrello, S. Calderara, and R. Cucchiara, “Latent space autoregression for novelty detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 481–490, Long Beach, California, June 2019.
- [20] R. T. Ionescu, F. S. Khan, M. I. Georgescu, and L. Shao, “Object-centric auto-encoders and dummy anomalies for abnormal event detection in video,” in *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7842–7851, Long Beach, California, June 2019.
- [21] R. Xie, N. M. Jan, K. Hao, L. Chen, and B. Huang, “Supervised variational autoencoders for soft sensor modeling with missing data,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2820–2828, April 2020.
- [22] Z. Teed and J. Deng, “RAFT: recurrent all pairs field transforms for optical flow,” in *Proceedings of the European Conference on Computer Vision*, pp. 402–419, Glasgow, Scotland, August 2020.
- [23] R. Morais, V. Le, T. Tran, B. Saha, M. Mansour, and S. Venkatesh, “Learning regularity in skeleton trajectories for anomaly detection in videos,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 11996–12004, Long Beach, California, June 2019.
- [24] W. Liu, W. Luo, D. Lian, and S. Gao, “Future frame prediction for anomaly detection—a new baseline,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6536–6545, Salt Lake City, UT, USA, June 2018.
- [25] L. Weixin Li, V. Mahadevan, and N. Vasconcelos, “Anomaly detection and localization in crowded scenes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 18–32, 2014.
- [26] Y. Chang, Z. Tu, W. Xie, and J. Yuan, “Clustering driven deep autoencoder for video anomaly detection,” in *Proceedings of the European Conference on Computer Vision*, pp. 329–345, Springer, Glasgow, Scotland, August 2020.
- [27] L. Ruff, R. A. Vandermulen, N. Gornitz et al., “Deep one-class classification,” in *Proceedings of the 35 th International Conference on Machine Learning*, Stockholm, Sweden, July 2018.

Research Article

Improved Deep Neural Network for Cross-Media Visual Communication

Yubo Miao 

College of Furniture and Art Design, Central South University of Forestry and Technology, Changsha, Hunan 410000, China

Correspondence should be addressed to Yubo Miao; t20100701@csuft.edu.cn

Received 14 March 2022; Accepted 5 April 2022; Published 30 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yubo Miao. 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.

Cross-media visual communication is an extremely complex task. In order to solve the problem of segmentation of visual foreground and background, improve the accuracy of visual communication scene reconstruction, and complete the task of visual real-time communication. We propose an improved generative adversarial network. We take the generative adversarial network as the basis and add a combined codec package to the generator, while configuring the generator and discriminator as a cascade structure, preserving the feature upsampling and downsampling convolutional layers of visual scenes with different layers through correspondence. To classify features with different visual scene layers, we add a new auxiliary classifier based on convolutional neural networks. With the help of the auxiliary classifier, similar visual scenes with different feature layers have a more accurate recognition rate. In the experimental part, to better distinguish foreground and background in visual communication, we perform performance tests on foreground and background using separate datasets. The experimental results show that our method has good accuracy in both foreground and background in cross-media communication for real-time visual communication. In addition, we validate the efficiency of our method on Cityscapes, NoW, and Replica datasets, respectively, and experimentally demonstrate that our method performs better than traditional machine learning methods and outperforms deep learning methods of the same type.

1. Introduction

Influenced by COVID-19, the home office has become a new office situation, and cross-media visual communication technology has become a key technology for telecommuting. Cross-media visual communication technology is also more common in our life. Thanks to the maturity of 5G technology, current visual communication technologies such as teleconferencing, webcasting, and remote manipulation play an important role in long-distance communication. Nevertheless, current software technologies still have some shortcomings that lead to the loss of the physical connection between the camera and the display, hindering the widespread use of cross-media visual communication systems [1]. Researchers in the literature [2] have tried to solve this problem by upgrading the hardware, but the hardware is too bulky to be mass produced. Some researchers made minor software adjustments and added

human eye image correction algorithms to optimize visual communication [3–5].

A cross-media visual communication system often consists of a foreground and a background, with the foreground representing the communicator involved in the visual communication and the background representing scenes other than the communicator [6, 7]. A complete visual communication system is shown in Figure 1. The preliminary visual communication research mainly relies on the collaborative work with stereo cameras to complete the restoration between virtual images to real scenes, and this method requires a high camera scene coverage; in addition, the synchronization and real time between cameras are also required. In later visual communication research, researchers gradually started to use dynamic planning algorithms to achieve control over the matching between virtual scenes by parallax standard lines, as this method could not control the differences between adjacent parallaxes [8–10].

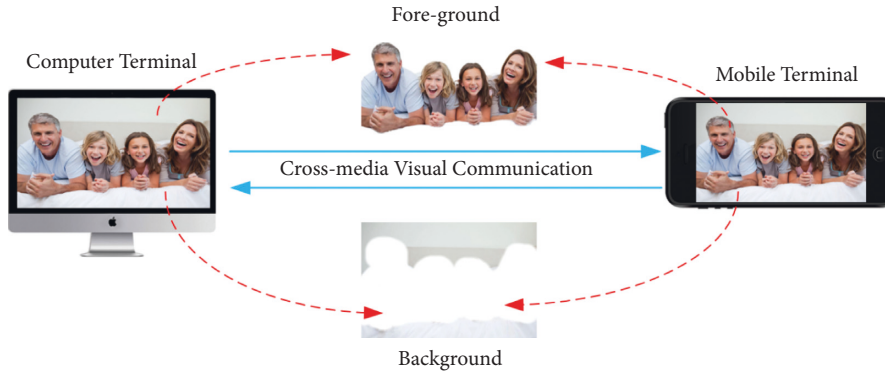


FIGURE 1: Cross-media visual communication background segmentation.

Later, researchers began to explore the rules from the hierarchical algorithm and take advantage of the hierarchical algorithm to exclude the smoothing constraint problem at the parallax end by hierarchical virtual matching layer by layer [11]. Adaptive optimization is also performed for poorer results at the boundaries, and each layer is considered as a uniform set, with each layer corresponding to a pixel depth offset, for parallax compensation. In the image recovery stage, the authors use motion estimation techniques to capture the motion of the virtual image and then match it to the displayed scene [12].

We experimentally verify the efficiency of traditional machine learning methods and deep learning methods in visual communication, and the deep learning methods perform better. So, in this study, we propose a visual communication method based on an improved generative adversarial network. The visual scene is layered by layer features, and foreground and background scenes are trained with unused methods and datasets, respectively. Then, we trained and verified the effectiveness of our method on a public dataset. For the scene segmentation problem of foreground and background, we added control experiments to verify the visual communication efficiency of foreground and background separately with specific datasets. The experimental trials demonstrate that our method is accurate and can achieve real-time performance in the visual communication of foreground and background.

The rest of this study is organized as follows. Section 2 introduces the research history and research results of cross-media visual communication. Section 3 describes in detail the principles and implementation processes related to the improved deep visual communication network. Section 4 shows the related experimental setup, the experimental dataset, and the analysis of the experimental results. Finally, Section 5 summarizes our research and reveals some further research work.

2. Related Work

Researchers in the literature [13] have proposed a simulated human eye image correction model to simulate the human eye's ability to process images. The model captures images through a single camera with its position and orientation

vectors and then maps the corrected images to the desired cross-media occasions for scene synthesis. In the literature [14], a pairwise polar-constrained view synthesis model is proposed to obtain a wider coverage of visual communication. The authors found in their experiments that the scene coverage of stereo cameras is wider than that of ordinary cameras, so the model uses two independent stereo cameras to capture the scene, and then, the captured model is transformed by 3D point features and mapped to three independent storage units, which can be called directly in scene reduction. Researchers in the literature [15] were inspired by experiments in scene recovery to propose a face model stereo matching algorithm, which has certain requirements for camera placement and requires a pair of simultaneous image captures at the same frame rate phase rate. The virtual camera is then used to synthesize the view, and the synthesized view is matched to the original view for stereo scene matching. Researchers in the literature [16] used a trinocular stereo camera for the first time. The images are captured through the synergy of cameras with different angles, and then, the scene is synthesized by the virtual view, and finally, the depth map is used to recover the communication signal that can support the 3D vision.

In addition to the optimization of the stereo camera matching algorithm, some researchers have discovered a new stereo matching technique, namely, the dynamic planning algorithm [17–19]. This algorithm determines the parallax by scan lines so that the virtual scene matching work can be performed in parallax standard lines. However, this algorithm cannot take into account the differences between adjacent scan lines well, which leads to the error between domains being expanded indefinitely. To address this problem, different dynamic planning algorithms are proposed in the literature [20] according to four different states, respectively. For example, in the case of occlusion, a dual-attention mechanism is invoked to complement the network structure. Besides, the authors also utilize a spatial filtering algorithm to smooth the model, which further improves the stability of the model. The literature [21] is optimized based on the former by adding two new states for segmenting foreground pixels and background pixels to obtain image segmentation capabilities with very high quality. The above research methods are

optimized from different perspectives, and all of them eventually improve the matching quality of virtual scenes.

For several years, Markov random fields have been gradually explored by researchers, and their excellent pixel-labeling ability has become a key player in dense stereo matching problems. Some researchers have found in their experiments that the Markov random field is gradually upgraded to a dedicated technique for graph cutting based on the Markov random field in the face of the transformation energy minimization problem, and the main principle of this technique is the confidence propagation minimization law [22–24]. We found in our experiments that the visual communication within and between scan lines is obtained from the Markov random field modeling so that excellent image segmentation capabilities can be obtained in the process of virtual image matching [25–28].

Researchers in the literature [29] found in their model optimization work that the continuity of the model’s scene smoothing variability and energy did not satisfy a normal distribution and hypothesized that the reason for the discrepancy was envisioned to be due to the difference in smoothing constraints at the parallax end. In the literature [30], to verify the former idea, a hierarchical stereo matching method was proposed to exclude the smoothing constraint problem at the parallax end layer by layer through hierarchical virtual matching. Also, adaptive optimization is performed for poorer results at the boundaries, and each layer is considered as a uniform set with each layer corresponding to a pixel depth offset for parallax compensation. In the image recovery stage, the authors employ motion estimation techniques to capture the motion of the virtual image and then match it with the displayed scene. The literature [31] innovatively proposed a Bayesian layering method, which concentrates the parallax between layers in a Bayesian decision framework, automatically traverses the parallax without layers by generalized expectation, finds the intermediate value, and finally compensates the parallax beyond the intermediate value. The literature [32] goes deeper in the study of layering, to preserve the boundaries between visual scene boundaries and to ensure the continuity of the boundaries. The authors proposed an iterative layering algorithm to achieve the optimal parallax range by successive iterations of the deep network, and this method somehow improves the real-time processing efficiency.

3. Method

3.1. Basic Structure. Generative adversarial networks have great applications in computer vision. Generative adversarial networks are essentially deep neural networks for generative models, which consist of two parts: a generator and a discriminator. The purpose of the network is to simulate learning training samples and generate samples with similar features to the training samples. During the training process, the generator generates a random fake sample from the input sample, and then, the discriminator directly determines the gap between the fake sample and the real sample, and if the gap is beyond the specified range, the fake sample

is returned, and the generator regenerates a new sample based on the returned information until the discriminator passes the sample. The structure of the generative adversarial network is shown in Figure 2.

We apply generative adversarial networks to cross-media visual communication. The accuracy of cross-media visual communication is improved by utilizing the mutual game learning approach of this network. Considering the scene complexity and image processing details of our visual communication, adaptive optimization based on generative adversarial networks will be considered to simulate optimal parallax for different visual communication levels. The details of the optimization will be explained in the subsequent subsections.

3.2. Generator. For generators, which are usually simple deconvolutional networks or fully connected neural networks, their main purpose is to generate dummy samples based on the output information, and their principle of action is shown in Figure 3. Finally, we usually analyze the generated output with features of different dimensions, and the feature dimensions are not strictly specified before training.

The literature [33, 34] propose an improved approach for generating adversarial networks, where the design of the generator as a matching encoder and decoder is very novel and effective. Therefore, we also adopt this combination of coding and decoding packages. We reconfigure the encoding and decoding parts in a cascade structure. The input data encoder, which consists of two convolutional layers, serves for downsampling, and then, I^{low} transforms different visual scene data features to the hidden layer. Inspired by the literature [35], we also used a structure of nine residual blocks to enrich different expression intensity features. The decoder consists of two deconvolution layers whose role is to be used for upsampling and to achieve the conversion of intensity features with the target expressions. All the above structures use a step size $s = 2$, kernel $= 3 \times 3$ setting, and all the convolutional layers are followed by a normalization operation and then activated nonlinearly by ReLU. For input surfaces with different expression midpoints, we use the X conv operator [36]. For a given K input points (p_1, p_2, \dots, p_k) , the K input points are weighted by a multilayer perceptron, and then, a $K \times K$ X transformation matrix, $X = MLP(p_1, p_2, \dots, p_k)$, is performed, followed by an element product summation, and is given to the X transform features in as a typical convolution operator treatment. Considering the influence of relative positions between different feature points, we define the X conv operator as follows:

$$F_p = X_{\text{conv}}(K, p, P, F), \quad (1)$$

$$X_{\text{Conv}}(K, p, P, F) = \text{Conv}(K, MLP(P - p) \times [MLP_{\delta}(P - p), F]), \quad (2)$$

where p represents the feature points, K represents the adaptive convolution kernel, $P = (p_1, p_2, \dots, p_k)^T$ represents

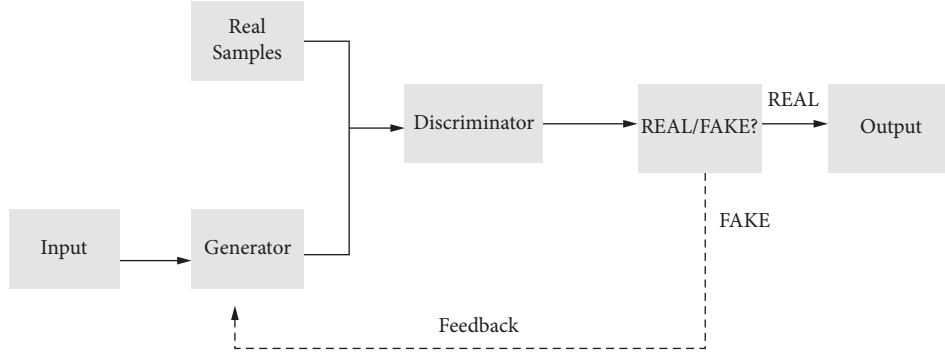


FIGURE 2: Generative adversarial network architecture.

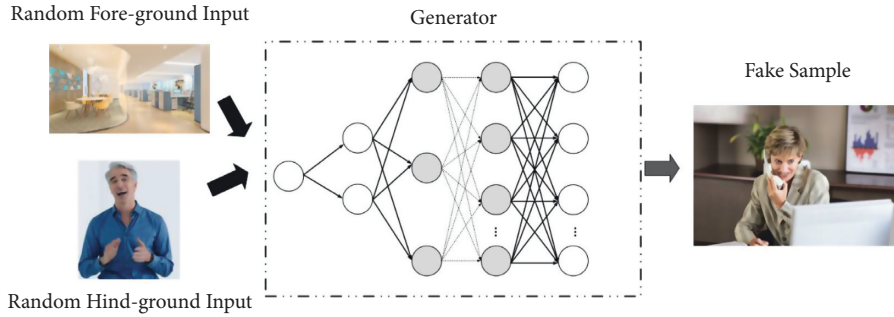


FIGURE 3: Cross-media visual communication generator process.

the K points in its neighborhood, and $F = (f_1, f_2, \dots, f_K)$ represents the features of different points. Using the principle of the X conv operator, we construct generators representing the parallax lines of visual scenes with different hierarchical structures, as shown in Figure 4. We apply the jump connection in the residual structure to the encoder and decoder to achieve the correspondence of random point coordinate information between the encoder and decoder.

3.3. Discriminator. In the structural design of the discriminator, we also use a combination of a deconvolutional and fully connected neural network designed to distinguish the difference between the fake and real samples generated by the generator. If the difference exceeds a predetermined value, and if it is trained by fine-tuning the discriminator's samples, it is fed back to the generator by backpropagation. The principle of its action is shown in Figure 5.

Low-level visual scene parallax is converted into high-level parallax. We use a generator to filter the underlying data density of high-level visual scene parallax features and a discriminator to separate the visual scene parallax features generated by the generator from the true high-level visual scene parallax features. Inspired by the study in the literature [37], we use alternating training patterns between the generator and discriminator to optimize the discretization problem between maxima and minima. For this purpose, we define the min-max problem as follows:

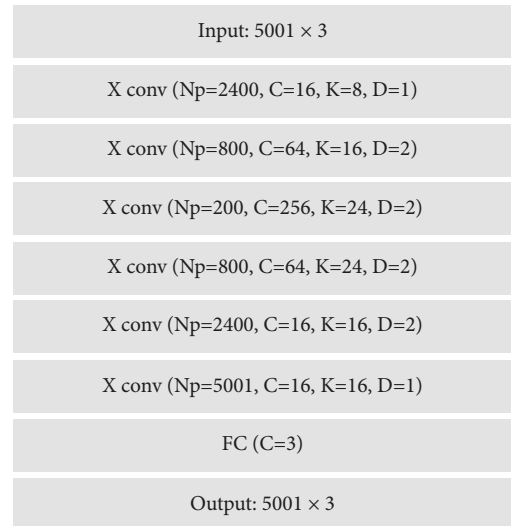


FIGURE 4: Detailed hierarchy of generators.

$$\min_{Gen} \max_{Dis} = E_I^{high} \log(Dis(I^{high})) + E_I^{low} \log(1 - Dis(Gen(I^{low}))), \quad (3)$$

where Gen represents the visual scene features generated by the generator, and Dis represents the visual scene features determined by the discriminator. $\{I^{low}, I^{high}\}$ represents a pair of parallax lines with different feature levels but the same visual scene class. The adversarial loss function equations for generator Gen and discriminator Dis are shown below.

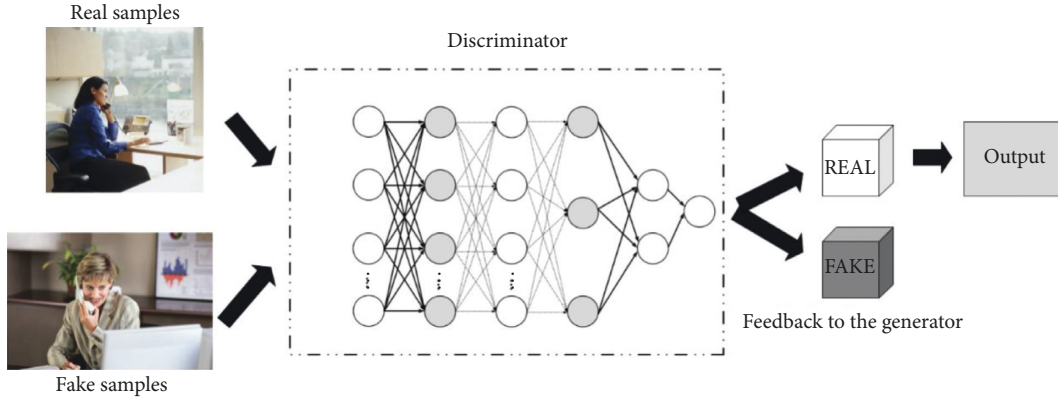


FIGURE 5: Cross-media visual communication discriminator process.

$$L_{G \rightarrow a \rightarrow dv} = -\frac{1}{N} \sum_{n=1}^N \log(\text{Dis}(\text{Gen}(I_n^{\text{low}}))), \quad (4)$$

$$L_{D \rightarrow a \rightarrow dv} = -\frac{1}{N} \sum_{n=1}^N \{ \log(\text{Dis}(I_n^{\text{high}})) + \log(1 - \text{Dis}(\text{Gen}(I_n^{\text{low}}))) \}, \quad (5)$$

where N denotes the total number of training samples. In the convergence process, the discriminator converges the features of the fake samples generated by the generator to the features of the real samples by setting the values so that the generator generates rich virtual visual scenes. Similarly, we constructed visual scene discriminators with different hierarchical structures. The hierarchical structure is shown in Figure 6.

3.4. Auxiliary Optimization. The scene recognition of different layers of features in similar representations is considered. We introduce auxiliary classifiers in generative adversarial networks for retaining feature information under different layers. In real life, each person is capable of generating hundreds of foregrounds and backgrounds with different meanings and hierarchies in different visual scenes. In our study, the generator is made to enable feature reconstruction of real visual scenes while keeping the visual scene category variables stable. We use the adversarial loss to rationally bootstrap the generator and train the auxiliary classifier to extract visual scene features. Perceptual loss is imposed when optimizing the generator and discriminator.

The auxiliary classifier can efficiently handle highly variable data. To match the adaptive linear fitting neural network, we introduce a convolutional neural network model. The dataset used for classifier training is processed real samples, all of which contain different layers of features. In our experiments, we build an online training method to guarantee its perfect convergence. In the subsequent network phase, we use a feature extraction layer instead of a fully connected layer, and the equation of the perceptual loss function in the feature extraction layer is shown below.

Input: 5001×3
X conv (Np=2400, C=16, K=8, D=1)
X conv (Np=800, C=64, K=16, D=2)
X conv (Np=200, C=256, K=24, D=2)
FC (C=1)
Mean
Output: 1×1

FIGURE 6: Detailed hierarchy of discriminators.

$$L_{\text{perceptual}} = \frac{1}{N} \sum_{n=1}^N \left\| \phi(G(I_n^{\text{low}})) - \phi(I_n^{\text{high}}) \right\|, \quad (6)$$

where ϕ denotes the feature extractor, and the biggest role of the perceptual loss function is to capture the variance distance between high-level sensitive features between similar visual scenes, thus maintaining the hierarchical stability of parallax features of visual scenes. In addition, it can also compensate the parallax lines of different layers of features to achieve a steady state.

To contribute to the consistency of the image representation, we use point-by-point loss optimization between the high-level scene representation feature I^{high} in the real sample and the synthetic virtual scene representation feature $\text{Gen}(I^{\text{low}})$ in the fake sample. As also mentioned in the literature [38], the distance between the intensity features is constrained by the L1 loss function, while L2 is more efficient in terms of integration performance. Their combined constraint equations are shown below.

$$L_{\text{pixel}} = \frac{1}{N_{\text{pixel}}} \sum_{i=1}^{N_{\text{pixel}}} \left\| \text{Gen}(I^{\text{low}})_i - I_i^{\text{high}} \right\|, \quad (7)$$

where N_{pixel} represents the geometric pixels in the two-dimensional sample and also denotes the three-dimensional sample point cloud data points. Combining the above loss

functions, the systematic loss function of our optimized generative adversarial network is formulated as follows:

$$L = \omega_1 L_{G_{-a} dv} + \omega_2 L_{pixel} + \omega_3 L_{perceptual}, \quad (8)$$

where ω_1 , ω_2 , and ω_3 represent the weighting coefficients. During the training process, we adopt an alternating training approach so that the generator repeatedly iterates to the optimum and finally generates the sample with the most similar physical signs to the real sample.

3.5. Improved Visual Communication Structure. To enhance the stability and accuracy of cross-media visual communication systems. We propose an improved deep neural network algorithm for visual communication, and the network underlying the method is a generative adversarial network. With our method, the stability of the cross-mediated visual communication system and the matching accuracy between communication messages can be further enhanced, and a visual scene feature grading model is also used to obtain the optimal parallax lines and compensation values. In our model, the generator synthesizes expressive features with high expression levels under the dual guidance of discriminators and auxiliary classifiers. For 2D samples, the convolutional neural network will assist in generating the samples, and for 3D data, they will be synthesized by the X conv operator. The features of different forms of virtual visual scenes are extracted and filtered in the joint output, and the real sample features of the auxiliary classifier are used as the guidance to fuse into a more comprehensive visual scene, so as to achieve an overfitting fit state with the real scene and achieve the maximum matching rate with the real scene. The detailed network structure is shown in Figure 7.

4. Experiment

4.1. Datasets. To validate the performance of our method, we selected Cityscapes, a public dataset for automatic segmentation of street scenes, NoW, a dataset for 3D face reconstruction, and Replica, a dataset for high-quality 3D reconstruction of indoor scenes, for experimental validation. Before the preprocessing operation on the above datasets, we normalize the visual scene boundaries for all datasets. In the preprocessing stage, the image and video frames are segmented into specified sizes, and different sizes of test methods correspond to different data. We randomly selected 80% of the samples as the training set and 20% as the test set. Next, we describe the three datasets in detail.

The dataset Cityscapes is a large-scale 3D video sequence of urban street scenes, divided into two types of strongly and weakly annotated frames. The biggest advantage of this dataset is its ability to automatically understand the semantics of the scenes and implement pixel-level semantic annotation of the full visual scenes. The main advantage of this dataset is that it can automatically understand the scene semantics and achieve full visual scene pixel-level semantic annotation. We mainly choose this dataset to

train the model for scene semantic segmentation performance. The dataset NoW is a face reconstruction dataset where all samples are independently performed by a 3D scanner. We mainly choose this dataset to obtain the matching accuracy of communicators in foreground reconstruction and to ensure a clear boundary between foreground and background. The dataset Replica is a high-quality indoor scene dataset, which contains 18 high realistic indoor visual scenes, and its biggest advantage is that all visual scenes are combined at the same level using plane mirrors and reflectors. We mainly choose this dataset to obtain the background matching accuracy and to achieve a high visual background reconstruction rate. The detailed information of the above dataset is shown in Table 1.

4.2. Experimental Results. To demonstrate the high efficiency of our method, we compared three methods, random forest (RF), RCNN, and BiLSTM. To ensure the accuracy of each method, each network is independently kept trained without the detection module during the training and tuning process. They are all trained using the ground truth of the three datasets mentioned in the previous section. The deep networks do not use any visual scenes, and among the evaluation metrics, we mainly look at precision (F), $F1$ score, and recall (R). The detection results of each metric are fed back to the individual metrics in the dataset, and the performance of text detection recognition can be indicated by the connection between the metrics, and the parallax baseline outputs after passing the visual scene classifier, which can also be labeled with the data to react to the accuracy of the virtual scene. There are two main steps in our cross-media visual communication work; first, we perform the visual scene classification work, and then, we perform the virtual scene reconstruction work. The visual scene classification performance is shown in Table 2.

Table 2 shows the visual communication efficiency of our method. The experimental results show that traditional machine learning algorithms poorly perform in visual communication. For example, the random forest algorithm is chosen to represent the traditional machine learning methods in this study, and its visual communication accuracy only reaches 62% at the highest. Among the methods in the deep learning category, RCNN is a better algorithm for visual processing, but its visual scene reconstruction accuracy is not as good as that of the BiLSTM algorithm. This is mainly due to the advantages brought by the special network structure of BiLSTM, which starts from the localization of the real visual scene and the virtual visual scene, and the parallax line compensation from the localization to the whole can optimize the error between adjacent domains well. Second, there is a small-to-large pattern that can expand the visual range and maximize the memory information fusion, which plays a key role in the process of matching the virtual visual scene with the real visual scene. Our method is based on generative adversarial networks, and then, a two-channel model is used to further obtain the bidirectional memory information of the adversarial depth

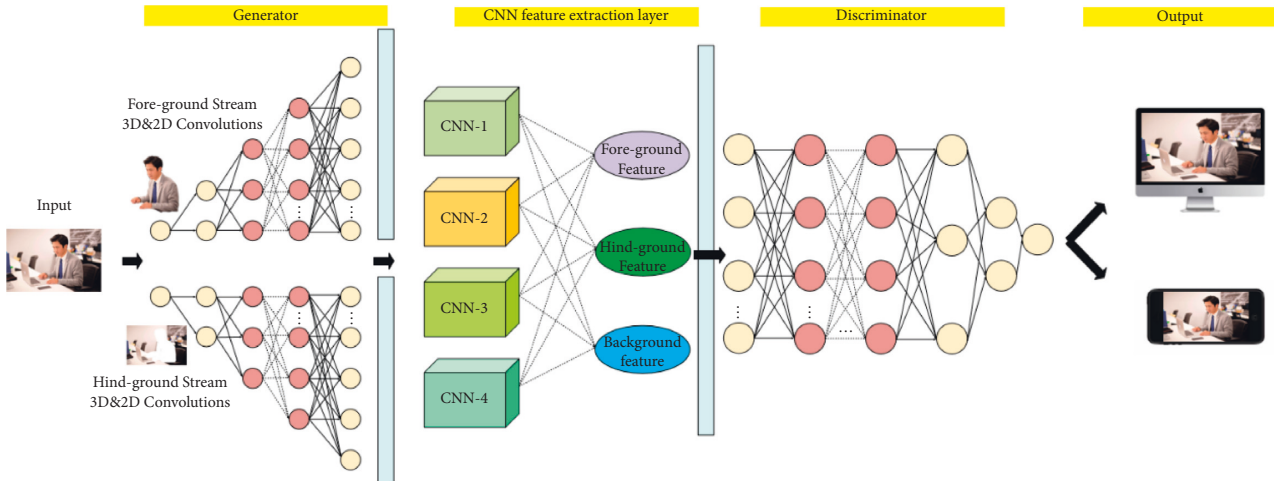


FIGURE 7: The structure of improved visual communication structure.

TABLE 1: The detailed information of datasets.

	Datasets		
	Cityscapes	NoW	Replica
Train	2301	3025	2361
Test	756	498	530
Total	3057	3523	2891

TABLE 3: Differences in visual communication between different datasets.

	Cityscapes	NoW	Replica
PER	0.61	0.87	0.92
VSRR	0.69	0.80	0.86
RSDR	0.76	0.88	0.84

TABLE 2: Results of visual communication by different methods.

	Cityscapes			NoW			Replica		
	P	R	F1	P	R	F1	P	R	F1
RF	0.58	0.56	0.61	0.62	0.66	0.62	0.62	0.55	0.63
RCNN	0.70	0.76	0.70	0.73	0.59	0.66	0.73	0.63	0.65
BiLSTM	0.78	0.83	0.83	0.77	0.73	0.74	0.85	0.79	0.79
Ours	0.91	0.93	0.86	0.89	0.85	0.86	0.90	0.92	0.90

network, which makes the visual communication more accurate.

After the visual scene step, it is difficult to perform the second stage of virtual scene reconstruction work because the traditional machine learning method random forest is not accurate enough in cross-media visual communication. Therefore, in the subsequent experiments, we excluded the traditional machine learning methods and kept only the deep learning methods as the efficiency comparison for visual communication. Before the virtual scene reconstruction work started, we compared three datasets to filter out the most suitable public dataset for cross-mediated visual communication because the data volume was too large. We evaluated the datasets in terms of parallax error rate (PER), virtual scene reconstruction rate (VSRR), and real scene distribution rate (RSDR), and the results of the dataset evaluation are shown in Table 3.

Table 3 shows the performance results of the street scene dataset Cityscapes, the face 3D reconstruction dataset NoW, and the indoor scene reconstruction dataset Replica in our own set of experimental metrics. From the

experimental results, we can see that the parallax error rate of street scenes is only 61%, which is a huge difference compared with other datasets. For this problem, we also did the subsequent experimental analysis, and because the complexity of the city street scene is much higher than other datasets, there will be frame loss and feature omission in the scene semantic reconstruction, which indirectly leads to the large error in the parallax of this dataset. At the same time, the Cityscapes data still poorly perform in the virtual scene reconstruction, which is also affected by the complexity of the city street scenes. Given that our cross-media visual communication research is more oriented toward online communication, the scene complexity of the Cityscapes dataset is given. In the subsequent experiments, we will select the NoW dataset and the Replica dataset for experimental validation. Among them, the NoW dataset targets foreground communicators for the scenario experiments. The Replica dataset targets the posterior scenes to verify the matching rate between the reconstructed virtual scenes and the real scenes.

To separate the foreground and background of the cross-media visual communication and highlight the detailed features of the characters, we re-edited the NoW dataset by reconstructing the 3D face reconstruction data in the dataset according to a standardized template and then took a smoothing process for each sample frame. In the scene reconstruction of foreground communicators, we mainly examine the metrics of face matching rate (FMR), foreground segmentation rate (FSR), and pixel offset rate (POR). The experimental results are shown in Table 4.

TABLE 4: Experiment results of foreground dataset.

	NoW		Fps
	Train	Test	
FMR	0.90	0.91	55
FSR	0.88	0.91	
POR	0.05	0.03	

TABLE 5: Experiment results of background dataset.

	Replica			Fps
	A (0–30%)	B (30%–70%)	C (70%–100%)	
BMR	0.96	0.90	0.83	56
HSR	0.97	0.88	0.80	
POR	0.02	0.05	0.09	

We also re-edited the Replica dataset and divided the indoor 3D scenes in the dataset into three levels, A, B, and C, where A denotes complexity less than 30%, B denotes complexity between 30% and 70%, and C denotes complexity greater than 70%. Then, a smoothing process is taken for each sample frame. In the scene reconstruction of the background, we mainly examine the metrics of background matching rate (BMR), background segmentation rate (HSR), and pixel offset rate (POR). The experimental results are shown in Table 5.

From the experimental results in Tables 4 and 5, it can be obtained that our method, in the foreground communicator scene reconstruction, has a cross-media visual communication accuracy of 91% and a pixel offset rate of only 5%. In the virtual visual scene reconstruction and real scene matching in the back view, the cross-media visual communication rate reaches over 80% in the most complex indoor scene, and the pixel offset rate is up to 9%, which satisfies the error controllable range. Meanwhile, the data transmission rate is between 55 fps and 56 fps, and the cross-media visual transmission of both foreground and background meets the real-time requirements of visual communication.

5. Conclusions

To improve the accuracy and real-time performance of cross-media communication, we propose an improved generative adversarial network. Considering the cascade structure of the deep adversarial network will be useful for cross-mediated visual communication. Therefore, we choose the generative adversarial network as the basis and add a combined codec packet to the generator, while configuring the generator and discriminator as a cascade structure, preserving the feature upsampling and downsampling convolutional layers of visual scenes with different layers through correspondence. To classify features with different layers of visual scenes, we added a new auxiliary classifier based on convolutional neural networks. With the help of the auxiliary classifier, similar visual scenes with different feature layers have more accurate recognition rates. In the experimental part, we validate the efficiency of our

method on Cityscapes, NoW, and Replica datasets, respectively, and experimentally demonstrate that our method performs better than traditional machine learning methods and outperforms deep learning methods of the same type. To better distinguish between foreground and background in visual communication, we perform performance tests on foreground and background with separate datasets. The experimental results demonstrate that our method has excellent accuracy in both the foreground and background of cross-media communication and can achieve real-time visual communication.

Cross-media visual communication is an extremely complex task that encompasses scene reconstruction, scene matching, and foreground-background segmentation. Each application scene has thousands of foreground-background combinations. In this study, we choose the simpler indoor scenes as our research point. However, for real cross-media visual communication scenes will be more complex. For complex visual scenes, our approach still poorly performs. In the next study, we will try to use the local perceptual layer of the BiLSTM network as the implicit layer visual scene extraction point to increase the visual scene capture range and improve the coverage of the model for visual scene types.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] L. Mhlbach, B. Kellner, and A. Prussog, "The importance of eye contact in a videotelephone service," in *Proceedings of the 11th International Symposium on Human Factors in Telecommunications*, France, September 1985.
- [2] R. Kollarits, C. Woodworth, and J. Ribera, "34.4: An eye contact camera/display system for videophone applications using a conventional direct-view LCD," in *Proceedings of the Society for Information Display, International Symposium*, pp. 765–768, San Diego, CA, USA, May 1996.
- [3] M. Ouyang, "Research on visual communication design system based on embedded network," *Microprocessors and Microsystems*, vol. 81, p. 103789, 2021.
- [4] G. S. Dhanesh and N. Rahman, "Visual communication and public relations: visual frame building strategies in war and conflict stories," *Public Relations Review*, vol. 47, no. 1, p. 102003, 2021.
- [5] V. Romero and A. Paxton, "Visual information and communication context as modulators of interpersonal coordination in face-to-face and videoconference-based interactions," *Acta Psychologica*, vol. 221, p. 103453, 2021.
- [6] G. Pressgrove, M. Janoske, and M. J. Haught, "Editors' letter: new research and opportunities in public relations and visual communication," *Public Relations Review*, vol. 44, no. 3, pp. 317–320, 2018.
- [7] M. Ståhl and H. Kaihovirta, "Exploring visual communication and competencies through interaction with images in social media[J]," *Learning, Culture and Social Interaction*, vol. 21, pp. 250–266, 2019.

- [8] K. Raoufi, C. Taylor, L. Laurin, and K. R. Haapala, "Visual communication methods and tools for sustainability performance assessment: linking academic and industry perspectives," *Procedia CIRP*, vol. 80, pp. 215–220, 2019.
- [9] D. Chai and Q. Peng, "Bilayer representation for three dimensional visual communication," *Journal of Visual Communication and Image Representation*, vol. 20, no. 8, pp. 552–562, 2009.
- [10] D. Liu, "Research on the analysis method of digital media art communication based on 3D image recognition," *Displays*, vol. 72, p. 102149, 2022.
- [11] A. S. Al-Subhi, "Metadiscourse in online advertising: e," *Journal of Pragmatics*, vol. 187, pp. 24–40, 2022.
- [12] A. Mikhailitchenko, R. G. Javalgi, G. Mikhailitchenko, and M. Laroche, "Cross-cultural advertising communication: visual imagery, brand familiarity, and brand recall," *Journal of Business Research*, vol. 62, no. 10, pp. 931–938, 2009.
- [13] J. Gemmell, K. Toyama, C. L. Zitnick, T. Kang, and S. Seitz, "Gaze awareness for video-conferencing: a software approach [J]," *IEEE MultiMedia*, vol. 7, no. 4, pp. 26–35, 2000.
- [14] R. Ruigang Yang and Z. Zhengyou Zhang, "Eye gaze correction with stereovision for video-teleconferencing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 7, pp. 956–960, 2004.
- [15] M. Ott, J. P. Lewis, and I. Cox, "Teleconferencing eye contact using a virtual camera," in *Proceedings of the INTERACT'93 and CHI'93 conference companion on Human factors in computing systems*, pp. 109–110, Amsterdam, Netherlands, April 1993.
- [16] J. Liu, I. Beldie, and M. Wopking, "A computational approach to establish eye-contact in video communication," in *Proceedings of the International Workshop on Stereoscopic and Three Dimensional Imaging (IWS3DI)*, pp. 229–234, Santorini, Greece, September 1995.
- [17] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, no. 2, pp. 139–154, 1985.
- [18] P. N. Belhumeur and D. Mumford, "A Bayesian treatment of the stereo correspondence problem using half-occluded regions," in *Proceedings of the 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 512, Champaign, IL, USA, 15–18 June 1992.
- [19] P. N. Belhumeur, "A Bayesian approach to binocular stereopsis," *International Journal of Computer Vision*, vol. 19, no. 3, pp. 237–260, 1996.
- [20] A. Criminisi, A. Blake, C. Rother, J. Shotton, and P. H. S. Torr, "Efficient dense stereo with occlusions for new view-synthesis by four-state dynamic programming[J]," *International Journal of Computer Vision*, vol. 71, no. 1, pp. 89–110, 2007.
- [21] V. Kolmogorov, A. Criminisi, and A. Blake, "Probabilistic fusion of stereo with color and contrast for bilayer segmentation[J]," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1480–1492, 2006.
- [22] Y. Boykov, O. Veksler, and R. Zabih, "Markov random fields with efficient approximations," in *Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231)*, pp. 648–655, Santa Barbara, CA, USA, June 1998.
- [23] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts[J]," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [24] J. Sun, H. Y. Shum, and N. N. Zheng, "Stereo matching using belief propagation," in *Proceedings of the European Conference on Computer Vision*, pp. 510–524, Copenhagen, Denmark, May 2002.
- [25] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]," *International Journal of Computer Vision*, vol. 47, no. 1, pp. 7–42, 2002.
- [26] M. F. Tappen and W. T. Freeman, "Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters," *IEEE Computer Society*, vol. 3, p. 900, 2003.
- [27] R. Szeliski, R. Zabih, and D. Scharstein, "A comparative study of energy minimization methods for Markov random fields with smoothness-based priors[J]," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 1068–1080, 2008.
- [28] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision[J]," *International Journal of Computer Vision*, vol. 70, no. 1, pp. 41–54, 2006.
- [29] D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities[J]," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 4, pp. 413–424, 1986.
- [30] S. Baker, R. Szeliski, and P. Anandan, "A layered approach to stereo reconstruction," in *Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231)*, pp. 434–441, Santa Barbara, CA, USA, 25–25 June 1998.
- [31] P. H. S. Torr, R. Szeliski, and P. Anandan, "An integrated Bayesian approach to layer extraction from image sequences [J]," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 297–303, 2001.
- [32] M. H. Lin and C. Tomasi, "Surfaces with occlusions from layered stereo," in *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Madison, WI, USA, 18–20 June 2003.
- [33] Y. H. Lai and S. H. Lai, "Emotion-preserving representation learning via generative adversarial network for multi-view facial expression recognition," in *Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pp. 263–270, IEEE, Xi'an, China, 2018.
- [34] H. Yang, U. Ciftci, and L. Yin, "Facial expression recognition by de-expression residue learning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2168–2177, Salt Lake City, UT, USA, 18–23 June 2018.
- [35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, Las Vegas, NV, USA, 27–30 June 2016.
- [36] Y. Li, R. Bu, and M. Sun, "Pointcnn: convolution on x-transformed points[J]," *Advances in Neural Information Processing Systems*, vol. 5, p. 31, 2018.
- [37] I. Goodfellow, J. Pouget-Abadie, and M. Mirza, "Generative adversarial nets[J]," *Advances in Neural Information Processing Systems*, vol. 27, 2014.
- [38] R. Huang, S. Zhang, T. Li, and H. Ran, "Beyond face rotation: global and local perception gan for photorealistic and identity preserving frontal view synthesis," in *Proceedings of the IEEE international conference on computer vision*, pp. 2439–2448, Venice, Italy, December 2017.

Review Article

Hyperspectral Image Classification: Potentials, Challenges, and Future Directions

Debaleena Datta ¹, **Pradeep Kumar Mallick** ¹, **Akash Kumar Bhoi** ^{2,3,4},
Muhammad Fazal Ijaz ⁵, **Jana Shafi** ⁶, and **Jaeyoung Choi** ⁷

¹*School of Computer Engineering, Kalinga Institute of Industrial Technology, Deemed to be University, Bhubaneswar 751024, India*

²*KIET Group of Institutions, Delhi-NCR, Ghaziabad-201206, India*

³*Directorate of Research, Sikkim Manipal University, Gangtok 737102, Sikkim, India*

⁴*AB-Tech eResearch (ABTeR), Sambalpur, Burla 768018, India*

⁵*Department of Intelligent Mechatronics Engineering, Sejong University, Seoul 05006, Republic of Korea*

⁶*Department of Computer Science, College of Arts and Science, Prince Sattam Bin Abdul Aziz University, Wadi Ad-Dawasir 11991, Saudi Arabia*

⁷*School of Computing, Gachon University, Seongnam-si 13120, Republic of Korea*

Correspondence should be addressed to Muhammad Fazal Ijaz; fazal@sejong.ac.kr and Jaeyoung Choi; jychoi19@gachon.ac.kr

Received 4 March 2022; Revised 22 March 2022; Accepted 30 March 2022; Published 28 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Debaleena Datta 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.

Recent imaging science and technology discoveries have considered hyperspectral imagery and remote sensing. The current intelligent technologies, such as support vector machines, sparse representations, active learning, extreme learning machines, transfer learning, and deep learning, are typically based on the learning of the machines. These techniques enrich the processing of such three-dimensional, multiple bands, and high-resolution images with their precision and fidelity. This article presents an extensive survey depicting machine-dependent technologies' contributions and deep learning on landcover classification based on hyperspectral images. The objective of this study is three-fold. First, after reading a large pool of Web of Science (WoS), Scopus, SCI, and SCIE-indexed and SCIE-related articles, we provide a novel approach for review work that is entirely systematic and aids in the inspiration of finding research gaps and developing embedded questions. Second, we emphasize contemporary advances in machine learning (ML) methods for identifying hyperspectral images, with a brief, organized overview and a thorough assessment of the literature involved. Finally, we draw the conclusions to assist researchers in expanding their understanding of the relationship between machine learning and hyperspectral images for future research.

1. Introduction

Hyperspectral imagery is one of the most significant discoveries in remote sensing imaging sciences and technological advancements. Hyperspectral imagery (HSI) is the technology that depicts the perfect combination of Geographic Information System (GIS) and remote sensing. Besides, HSI has several advantages such as ecological protection, security, agriculture and horticulture applications, crop specification and monitoring, medical diagnosis,

identification, and quantification [1]. RGB images are made up of three dimensions: width, height, and 3 color bands or channels consisting of color information, that is, red, green, and blue. They are stored as a 3D byte array that explicitly holds a color value for each pixel in the image; a combination of RGB intensities put down onto a color plane. However, in contrast, HSI comprises thousands of hypercubes and hence possesses a large resolution and an enormous amount of embedded information of all kinds—spectral, spatial, and temporal. This information enables various applications to

detect and characterize land covers, which are most significantly explored [2]. RGB images are captured by digital RGB cameras capable of characterizing objects only based on their shape and color. Moreover, the embedded information is minimal since only three visible bands are available in the human visibility range. The HSI, on the other hand, is captured by specialized airborne hyperspectral sensors placed on artificial satellites, that is, spectrometers. They have a broad range of scenes by acquiring large numbers of consecutive bands, not confined to the visible light spectrum and through a wider spectral band-pass. However, compared to the digital sensor that absorbs light in just three wide channels, a hyperspectral sensor's channel width is much narrower, making the spectral resolution and data volume much higher, resulting in hurdles to store, mine, and manage [3]. Furthermore, processing these data with a massive number of bands imposes many obstacles such as noise-causing image calibration, geometric distortion, noisy labels, and limited or unbalanced labeled training samples [4–6], that is, Hughes phenomenon and dimensionality reduction-related artifacts: overfitting, redundancy, spectral variability, loss of significant features between the channels, etc. [7].

Classifying HSIs is considered to be an intrinsically nonlinear problem [8], and the initial approach by linear-transformation-based statistical techniques such as principle component analytical methods, that is, principal component analysis (PCA) [9] and independent component analysis (ICA) [10]; the discriminant analytical methods, that is, linear [11] and fisher [12]; wavelet transforms [13]; and composite [14], probabilistic [15], and generalized [16] kernel methods, had shown promising outcomes. Still, their focus was limited to spatial information. They emphasized that the feature extractor techniques assisted by some basic random classifiers that lead to complexity in terms of cost, space, and time are not sufficiently accurate. After the success of these traditional methodical techniques assigned for HSI classification, researchers became keenly interested in applying the most recent emerging but not tedious computer-based methods that made the entire process smoother and vicinal to perfection. Study advancements suggest that the last decade can be considered the most escalating era regarding computer-based technologies due to the emergence of machine learning (ML). ML is an algorithmic and powerful tool that resembles the human brain's cognition. It simply represents a complex system by holding abstraction. Hence, it can reduce complexities and peep into the insights of the vast amount of HS data to fetch out the hidden discriminative features, both spectral and spatial [17]. Thus, it overcomes all the stumbling blocks to achieve the desired accuracy in identifying the classes that the objects of the target HSI data belong to. Hence, they act as all-in-one techniques that can serve the purpose without further assistance. Keeping this in mind, we conducted an extensive survey based on the various discriminative machine and deep learning (ML, DL) models for HSI. In most of the literature studies, the HSI datasets that are commonly used for landcover classification are AVIRIS Indian Pines (IP), Kennedy Space Center (KSC), Salinas Valley (SV), and

ROSIS-03 University of Pavia (UP), along with less frequently used Pavia Center, Botswana, University of Houston (HU), etc. They are pre-refined and made publicly available on [18] for download and perform operations.

The motivation of our work is divided into three parts. First, a novel methodology is proposed for the review work that is entirely systematic and helps find the inspiration in forming the research gaps and embedded questions after going through a large pool of research articles. Second, this work focuses on the current advancements of ML technologies for classifying HSI, with their brief, methodical description and a detailed review of the literature involved with them. Finally, the inferences are drawn and help the researchers boost knowledge for their future research. The key contributions made to the research field on hyperspectral imagery by our novel effort are as follows:

- (1) The thorough revision of the analytical and classification work carried out to date on HS imagery by employing ML/DL techniques.
- (2) Emphasis on the categorized methods explored and practiced so far in an overly frequent manner. Also, it includes a brief interpretation of the most recent technologies and the highlighted hybrid techniques.
- (3) An open knowledge base that acts as a reservoir of relevant information that is listed out that interprets all research on each mentioned technique in terms of their methodology, convenience and limitations, and future strategies. This illustration might administrate in making a proper choice of objective for further research on the field of HSIs.
- (4) Explicit idea of the growth of interest in the concerned field that would attract researchers to invest themselves with a coherent, substantial specification (benefaction and drawbacks) of all the methods, individually, that contributes academically to the researchers about their favorable result and the difficulties for a chosen technique.
- (5) A transitory rendition of the most recent research on HSIs signifies the currently adapted technologies as hot spots. Also, focus on the research areas about the interest that could apply to others, that is, the hybridized methods popular among researchers to address the problem and achieve the desired experimental results.

The rest of the article is arranged as follows: Section 2 briefly explains the constraints faced by the researchers in dealing with HSI; Section 3 represents the methodology for the research along with the motive behind this review; Section 4 describes seven ML techniques, namely, support vector machine (SVM), sparse representation (SR), Markov random field (MRF), extreme learning machine (ELM), active learning (AL), deep learning (DL), and transfer learning (TL); Section 5 shows up the complete summary of the literature review work in the form of answers to the research questions; Section 6 depicts the conclusions; and Section 7 explains the limitations and future work.

2. Constraints of HSI Classification

Since their emergence, several difficulties have caused issues in analyzing and performing operations on hyperspectral images. Initially, it suffered from spectroscopy technology due to the bad quality of hyperspectral sensors and poor quality with insufficient data. However, along with the advancement in applied science, things have come to ease, but there are still some well-known nondispersible hitches that need to be overcome. Some of them are stated as follows:

- (a) Lack of high-resolution Earth observation (EO) noiseless images: During the initial stage of the discovery of spectrometers, they were not very efficient. Due to this, noises caused by water vapor, atmospheric pollutants, and other atmospheric perturbations modify the signals coming from the Earth's surface for Earth observations. Several efforts have been made over the last decades to produce high-quality hyperspectral data for Earth observation and develop a wide range of high-performance spectrometers that combines the power of digital imaging, spectroscopy, and extracting numerous embedded spatial-spectral features [19].
- (b) Hindrances in the extraction of features: During data gathering, redundancy across contiguous spectral bands results in the availability of duplicated information, both spatially and spectrally, obstructing the optimal and discriminative retrieval of spatial-spectral characteristics [7].
- (c) The large spatial variability and interclass similarity: The hyperspectral dataset collected contains unusable noisy bands due to mistakes in the acquisition that result in information loss in terms of the unique identity, that is, the spectral signatures and excessive intraclass variability. Furthermore, with poor resolution, each pixel comprises broad spatial regions on the Earth's surface, generating spectral signature mixing, contributing to the enhanced interclass similarity in border regions, thus creating inconsistencies and uncertainties for employed classification algorithms [19].
- (d) Limitation of available training samples and insufficient labeled data: Aerial spectrometers cover significantly smaller areas, so they can only collect a limited number of hyperspectral data. That leads to the restriction of the number of training samples for classification models [20]. In addition, HSIs typically contain classes that correspond to a single scene, and available classification models' learning procedures require labeled data. However, labeling each pixel requires human skill, which is arduous and time-consuming [21].
- (e) Lack of balance among interclass samples: The class imbalance problems, where each class sample has a wide range of occurrences, diminish the usefulness of many existing algorithms in terms of enhancing minority class accuracy without compromising

majority class accuracy, which is a difficult task in and of itself [22].

- (f) The higher dimensionality: Due to incorporating more information in multiple channels, such as high-band pictures increase estimation errors. The curse of dimensionality is a significant drawback for supervised classification algorithms, as it significantly impacts their performance and accuracy [23].

The possible solutions to the above limitations that also represent the possible operations that are performed to analyze and comprehend the HSIs can be (1) technological advancement to make versatile and robust hardware for the spectrometers to capture the scenes more accurately, (2) spectral unmixing and resolution enhancement for better feature extraction and distinguishing capability of the embedded objects, (3) image compression-restoration and dimensionality reduction for addressing the high-dimensions and lack of data, and (4) use of robust classifiers that are capable of dealing with the above issues as well as promote fast computation ability [7].

These hurdles were very prominent for the methods that classify HSI based on the feature extraction from HSI. After ML/DL came into the scene, the operations on HSI became effortless as explicit feature extraction is not needed, and it has also many advantages such as great dealing with noise and time complexity. However, ML/DL acquires a few drawbacks in specific criteria [19], including parameter-tuning and numerous local minima problems in training procedures and compression [20] overfitting, optimization, and convergence problems despite many positive aspects.

3. Research Methodology

This section is divided into three categories that will assist in understanding the review procedure and its ambition.

3.1. Planning of the Review. Three systematic advances are utilized that comprise the planning behind our work. First, based on efficacy and frequency of applicability on classifying HSIs, seven most recently used ML techniques have been chosen in this article for review, which establishes the operational relationship and compatibility with the issue of categorizing the land covers of a particular scene captured as HSI. Second, this relationship provides all the shortfalls and benefits of those methods and their potential possibilities. Finally, we identified the limitations of our present review work and how to rectify them in the future.

3.2. Conducting the Review. The entire review work has been conducted in the following steps:

- (a) Collection of literature: The literature studies have been collected based on the keywords: "Hyperspectral image classification," "Machine learning techniques," "Deep learning techniques," from the most relevant search engine, that is, Google (Google Scholar), which provides the scholarly articles for the concerned topic. These literature studies include

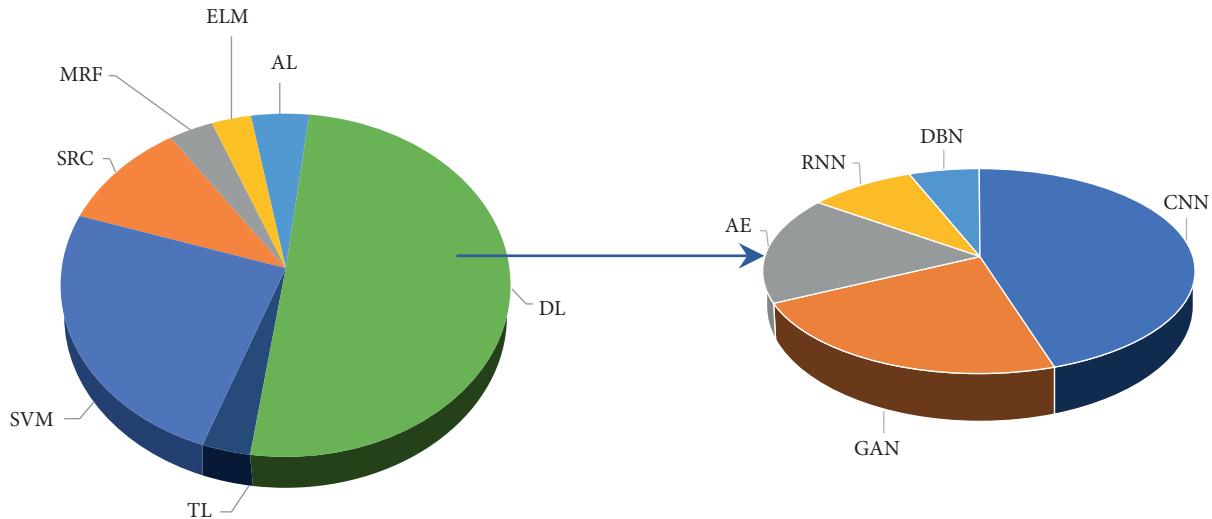


FIGURE 1: The statistical pie-charts of screened articles on ML/DL techniques used for HSI classification (source: SCI, SCIE, Scopus, WoS).

Web of Science (WoS), Scopus, SCI, and SCIE-indexed and SCIE-related articles, both journals and conferences. Several methods are utilized throughout the literature that assist the classification of hyperspectral data, out of which ML techniques seem to be more convenient and promising.

(b) Screening: The collected research papers depict raw data, sorted categorically according to the chronological order of the ML techniques used over the periods. The screening was accomplished based on the following constraints:

- (i) Time Period: The studies published in the range of 2010–2021 are included in this work. Studies published before 2010 are not included.
- (ii) Methodology: The studies on HSI’s analytical operations (denoising, spectral unmixing, etc.) other than classifying the underlying land covers are rejected.
- (iii) Type: The studies that deal with the hyperspectral images of a particular land scene are considered, discarding the medical hyperspectral imagery, water reservoir, etc.
- (iv) Design of study: The studies comprising experimental outcomes and the elaboration of the models are accepted; other literary-based articles or review papers are only for primary knowledge gain.
- (v) The language used: The studies written in the English language are only considered.

Figure 1 represents the total number of the literary studies screened individually on each of the categories of chosen ML techniques in the form of pie-charts with a percent-wise pattern. Figure 2 is a standard graphical depiction of the number of most recent articles that we screened for each chosen ML-based method in the period ranging from 2015 to 2021.

(c) Selection: Out of all the papers screened based on the abovementioned criteria, a few most eligible are handpicked. The selection has been made keeping specific parameters: the modeling strategy and algorithm and its suitability with the modern technological scenario. The final result is the corresponding overall accuracy (COA) for each dataset used, preferably journals with a good citation index.

(d) Analysis and inference: These selected papers are thoroughly reviewed to determine their contribution, restrictions, and future propositions. Based on this analysis, the deductions are drawn to show the pathway of further research.

3.3. *Research Investigations (RI)*. The analysis arises some of the queries:

RI 1: What is the significance of traditional ML and DL for analyzing HSI?

RI 2: How is ML/DL more impactful on HSI than other non-ML strategies?

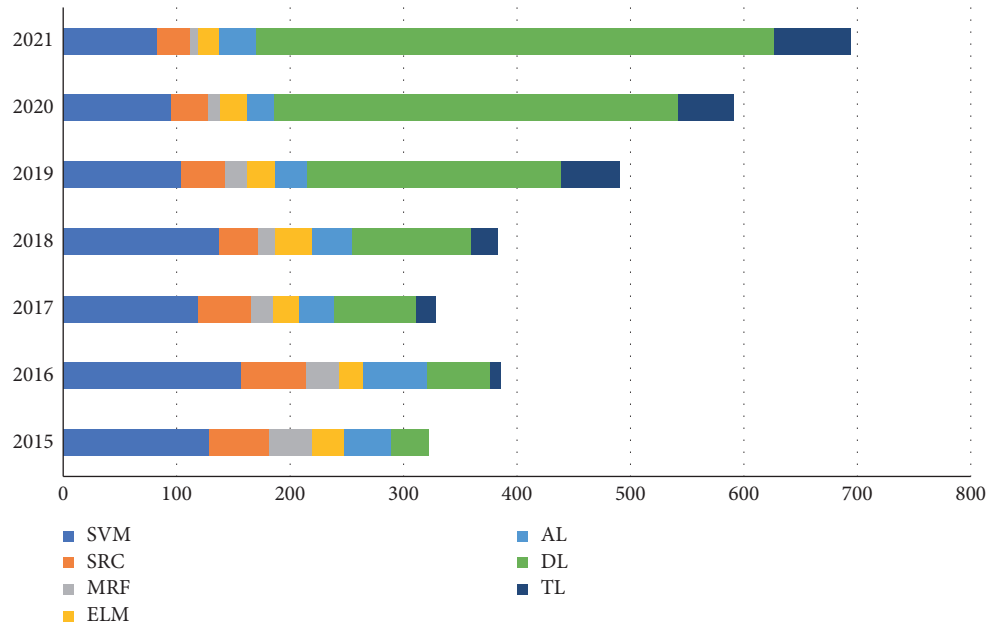
RI 3: What are the advantages and challenges faced by the researchers for the chosen ML/DL-based algorithm for HSI classification?

RI 4: What are the emerging literary works of ML/DL on HSI classification in the year 2021?

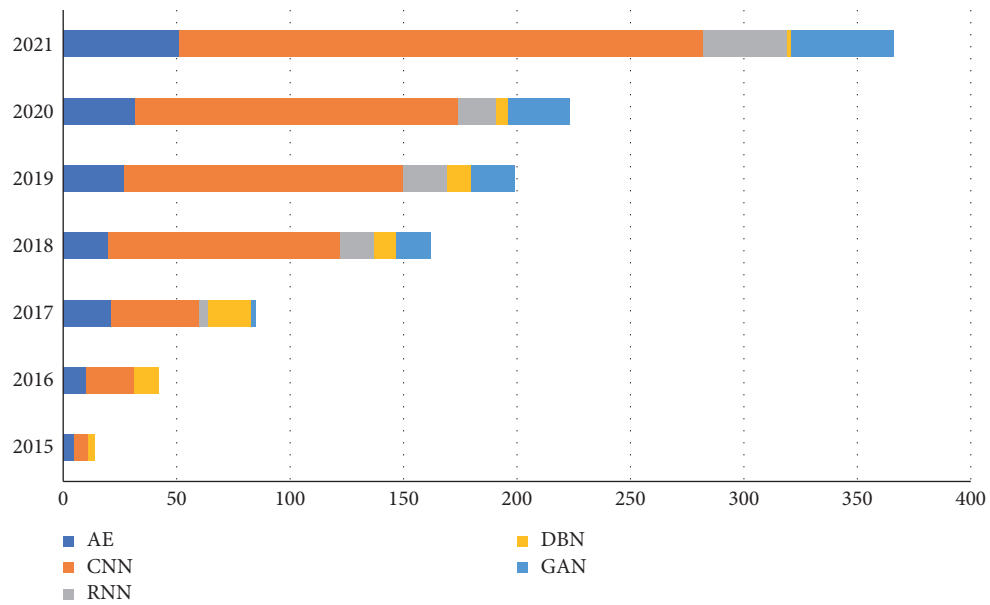
RI 5: How are ML- and DL-based hybrid techniques helping scientists in HSI classification?

RI 6: What are the latest emerging techniques associated with addressing classifying HSIs?

3.4. *Datasets*. The HSI datasets are pre-refined and made publicly available for download and perform operations. There are six datasets that are described here in a concise manner:



(a)



(b)

FIGURE 2: The statistical bar graph of screened articles on ML/DL techniques used for HSI classification from 2015 to 2021 (source: SCI, SCIE, Scopus, WoS): (a). ML. (b). DL.

(i) AVIRIS Indian Pines: This dataset was taken by airborne visible infrared imaging spectrometer (AVIRIS) sensor, on June 12, 1992. The scene captured here was Indian Pines test site in North-Western Indiana, USA, and contains an agricultural area exemplified by its crops of regular geometry and some irregular forest zones. It consists of 145×145 pixels with a spectral resolution of 10 nm and a spatial resolution of 20 mpp and 224 spectral reflectance bands in the wavelength range

$0.4\text{--}2.5 \mu\text{m}$, out of which 24 noisy bands are removed due to low signal-to-noise ratio. The scene contains 16 different classes of land covers.

(ii) Salinas Valley: This scene was obtained by AVIRIS sensor over various agricultural fields of Salinas valley, California, USA, in 1998. The scene is characterized by a high spatial resolution of 3.7 mpp and a spectral resolution of 10 nm. The area is covered by 512×217 spectral samples with a wavelength range of $0.4\text{--}2.5 \mu\text{m}$. Out of 224 reflector

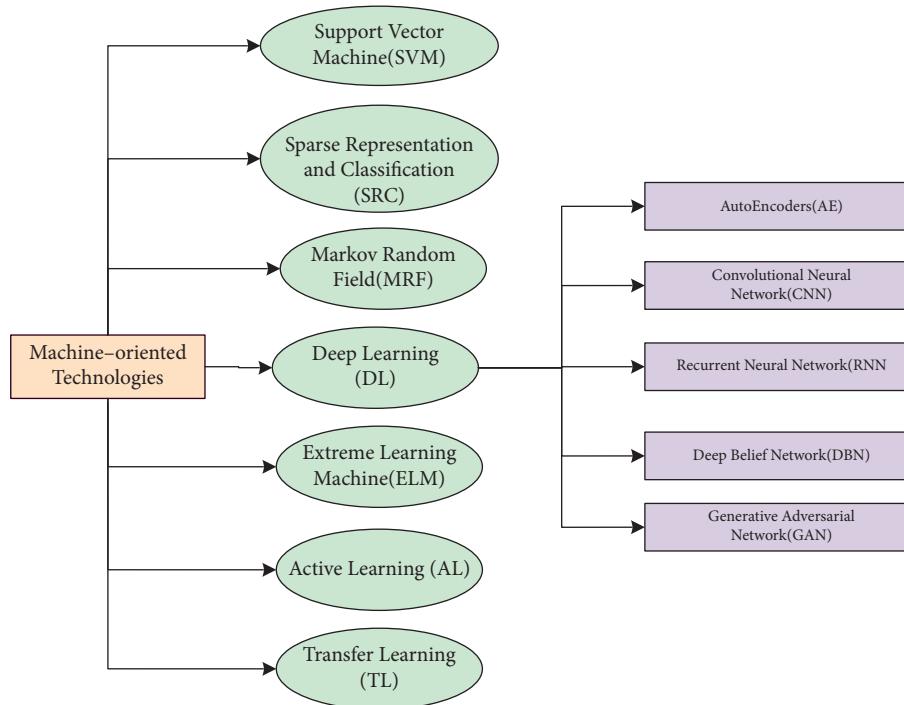


FIGURE 3: The categories of the eminent machine learning techniques used for HSI classification.

bands, 20 noisy bands are discarded due to water absorption coverage. The scene comprises 16 different land classes.

- (iii) Pavia Center: This scene was captured by a reflective optics system imaging spectrometer (ROSIS-03) sensor during a flight campaign over Pavia, northern Italy. It possesses 115 spectral bands, out of which only 102 are useful. Its spectral coverage is $0.43\text{--}0.86\ \mu\text{m}$, with a spectral resolution of 4 nm and a spatial resolution of 1.3 mpp defined by $1096 * 1096$ pixels. There are 9 different land cover classes in the area.
- (iv) Pavia University: This scene was also captured by the same sensor at the same time as Pavia center, over the University of Pavia in 2001. It has the same structural features as the Pavia center, only contrasting in considering 103 bands out of 115 bands with a size of $610 * 340$ are taken after discarding 12 noisy bands. The scene contains 9 classes with urban environmental constructions.
- (v) Kennedy Space Center: This scene was acquired by NASA AVIRIS sensor over Kennedy Space Center, Florida, USA, on March 23, 1996. It was taken from an altitude of approximately 20 kilometres, having a spatial resolution of 18 kilometres and a spectral resolution of 10 nm. The wavelength range of the scene is $0.4\text{--}2.5\ \mu\text{m}$ with the special size of $512 * 614$ pixels; 24 of 48 bands were removed for a low signal-to-noise ratio. The ground contains 13 predefined classes by the center personnel.
- (vi) Botswana: The scene was obtained by the Hyperion sensor placed on the NASA EO-1 satellite over

Okavango delta, Botswana, South Africa, on May 31, 2001. It has a special resolution of 30 metres and a spectral resolution of 10 nm while taken at an altitude of 7.7 kilometres. Out of 242 bands containing $1476 * 256$ pixels, with a wavelength range of $400\text{--}2500\ \text{nm}$, 97 bands are considered to be water-corrupted and noisy; hence, 145 remaining are useful. The scene comprises 14 land cover classes.

4. Machine Learning-Based Techniques for HSI Classification

ML technologies are not only intelligent and cognitive, but also their accuracy is skyrocketing due to their embedded mechanical abilities such as extraction, selection, and reduction of joint spatial-spectral features as well as contextual ones [24–26]. Moreover, the hidden dense layers with various allocated functions of the extensive networks work as intelligent learners by creating dictionaries or learning spaces to store deterministic information and then separate the landcover classes through its classification units [27–29]. The latest ML techniques that assist in classifying the hyperspectral data, that is, SVM, SRC, ELM, MRF, AL, DL, and TL, are shown categorically in Figure 3 and are discussed hereafter in detail.

4.1. Support Vector Machine (SVM). SVM is an innovative pattern-recognition technique rooted in the principle of statistical learning. The rudimentary concept of SVM-based training can unravel the ideal linear hyperplane so that the predicted classification error is mitigated, be it for binary or

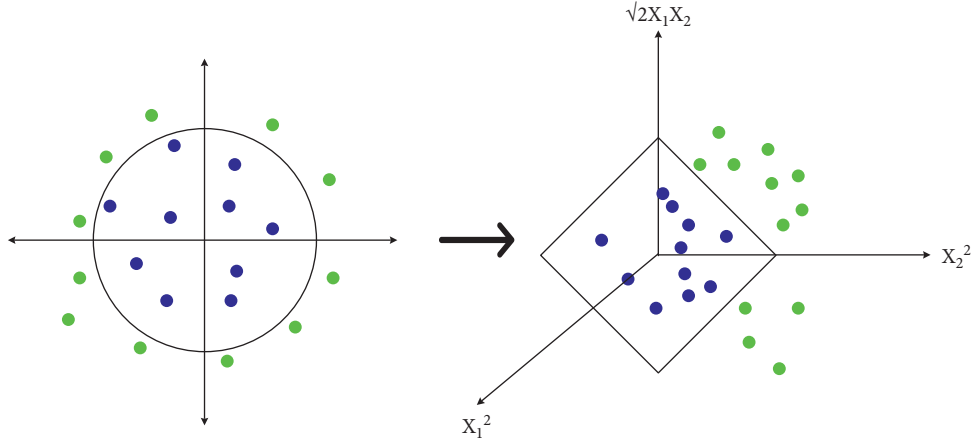


FIGURE 4: Classification strategy by multiclass SVM.

multiclass purposes [30], as depicted in Figure 4. For linearly separable binary classification, let (x_i, y_i) be the standard set of linearly separating samples with $x \in (\mathbb{R})^N$ and $y \in \{-1, +1\}$. The universal formula of linear decision hyperplane in n -dimensional space with the classification hyperplane is

$$g(x) = w^T \cdot x + b = 0, \quad (1)$$

where w is the weight directional vector and b is the slope of the hyperplane. A separating hyperplane with margin $2/\|w\|$ in the canonical form must satisfy the following constraints:

$$y_i [(w^T \cdot x_i) + b] \geq 1. \quad (2)$$

For multiclass scenarios, we presumably transform the datapoints to S , a probable infinite-dimensional space, by a mapping function ψ defined as $\psi(x) = (x_1^2, x_2^2, \sqrt{2}x_1x_2)$, $\mathbf{x} = (x_1, x_2)$. Linear operations performed in S resemble nonlinear processes in the original input space. Let $K(x_i, x_j) = \psi(x_i)^T \psi(x_j)$ be the kernel function, which remaps the inner products of the training dataset.

Constructing SVM requires values of the constants, that is, Lagrange's multipliers, $\alpha = (\alpha_1, \dots, \alpha_N)$ so that

$$P(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j). \quad (3)$$

is maximized with the constraints with respect to α :

$$\sum_{i=1}^N \alpha_i y_i = 0, \quad \alpha_i \geq 0 \text{ for all } \alpha_i. \quad (4)$$

Because most α_i are supposedly equal to zero, samples conforming to nonzero α_i are support vectors. Conferring to the support vectors, the modified optimally ideal classification function is

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b. \quad (5)$$

The application of SVM for classifying HSI started two decades ago [31, 32]. Focusing on the potentially critical issue of applying binary SVMs [33], fuzzy-based SVM [34] as

fuzzy input-fuzzy output support vector machine (F2-SVM), SVM evolved to dimensionality reduction and mixing of morphological details [35]. It also assisted particle swarm optimization (PSO) [36] and wavelet analysis with semi-parametric estimation [37], as the classifier "wavelet SVM" (WSVM). Table 1 summarizes the research carried out so far for the classification purpose of HSI using SVM.

4.2. Sparse Representation and Classification (SRC)

Sparse method depends on dictionary learning that enhances and rectifies the values of parameters based upon the current training observations while accumulating the knowledge of the previous observations prior. It then generates the sparse coefficient vector using sparse coding. This method is supremely efficient as it embeds dictionary learning to extract rich features embedded inside the HSI dataset. SR can classify images pixelwise by representing the patches around the pixel with a linear combination of several elements taken from the dictionary. The generalization of SRC called multiple SRC (mSRC) has three chief parameters—patch size, sparsity level, and dictionary size. Dictionary learning is the first step for sparse, using K-SVD algorithm. Let $Y = [y_1, y_2, \dots, y_N]$ be a matrix of L2-normalized training samples $y_i \in \mathbb{R}^m$ [45–47].

The size of patches around the pixel is

$$\min_{D, B} \|Y - DB\|_F^2 \text{ such that } \|b_i\|_0 \leq S, \text{ for all } i, \quad (6)$$

where D is a member of $\mathbb{R}^{m \times n}$ is the learned over a complete dictionary, with $n > m$ atoms, $B = [b_1, b_2, \dots, b_m]$ represents the matrix of corresponding sparse coding vectors $b_i \in \mathbb{R}^n$, and $\|\cdot\|_F$ is the Frobenius norm. Sparsity S limits the number of nonzero coefficients in each b_i . The next step sparse coding is provided with dictionary D and represents y as a linear combination of $y = D\hat{b}$ where \hat{b} is sparse. For the final classification step, suppose for each class $j \in \{1, \dots, M\}$ of an image, a dictionary D_j is trained. Then, the classification of a new patch y_{test} is achieved by estimating a representation error. The class assignments rule [47] is calculated through a pseudoprobability measure $P(C_j)$ for each class error E_j as

TABLE 1: Summary of review of HSI classification using SVM.

Year	Method used	Dataset and COA	Research remarks and future scope
2011	Multiclass SVM [38]	San Diego3—98.86%	Outperforms traditional SVM and deals better with Hugh's effect
2012	Fuzzy decision tree-support vector machine (FDT-SVM) [39]	Washington DC mall—94.35%	Efficient testing accuracy truncated computational and storage demand, understandable edifices, and reduction of Hugh's effect
2014	Semi-supervised SVM kernel-spectral fuzzy C-means (KSFCM) [40]	IP—98.52%	Enhanced classification and clustering by fully exploring both labeled and unlabeled samples
2014	SVM-radial basis function (SVM-RBF) [41]	IP—88.7%, UP—94.7%	Outperforms other existing kernel-based methods
2015	Regional kernel-based SVM (RK SVM) [42]	UP—95.40%, IP—92.55%	Outperforms pixel-point-based SVM-CK
2017	Multiscale segmentation of super-pixels (MSP-SVMsub) [43]	UP: MSP-SVMsub—97.57%, IP: MSP-SVMsub—95.28%	Solving classic OBIC-based methods with difficulties determining the appropriate segmentation size reduces the Hughes phenomenon
2018	Extended morphological profiles (EMP), differential morphological profiles (DMP), Gabor filtering with SVM [44]	UP: MFSVM-GF—98.46%, IP: MFSVM-GF—98.01%	Outruns several advanced classifiers: SVM, super-pixel-based SVM, SVM-CK, multifeature SVM, EPF
2019	SVM-PCA [24]	IP—91.37%, UP—98.46%	Outperforms Naïve Bayes, decision tree k-NN

$$j^* = \arg \max_j P(C_j), \quad \text{where, } P(C_j) = \frac{1}{M-1} \frac{\sum_{k=1, k \neq j}^M E_k}{\sum_{k=1}^M E_k}. \quad (7)$$

mSRC obtains residuals of disjoint sparse representation of y_{test} for all classes j . Each dictionary D_j is updated by eliminating nonzero atoms from \hat{b}_j after each of k iterations and y_{test} is assigned to the class, using Q total iterations:

$$D_j = \arg \max_j \sum_{k=1}^Q P_k(C_j). \quad (8)$$

Sparse representation is an essential and efficient machine-dependent method in many areas, including denoising, restoration, target identification, recognition, and monitoring. It may grow even more vital when associated with logistic regression, adaptivity, and super-pixels to extricate the joint features globally and locally. SR has a very high potential of being associated with methods such as PCA, ICA, Markov random fields, conditional random fields, extreme learning machines, and DL methods such as CNN and graphical convolutional network. Table 2 gives a summary of the research performed so far for the classification purpose of HSI employing SRC.

4.3. Markov Random Field (MRF). MRF describes a set of random variables satisfying Markov probability, depicted by undirected graphs. It is similar to the Bayesian network but, unlike it, undirected and cyclic. An MRF is represented as a graphical model of a joint probability distribution defined in Figure 5. The undirected graph of MRF, $G = (V, E)$, in which V is the nodes representing random variables.

Based on the Markov properties [57], the neighborhood set N_c of a node c is defined as

$$N_c = \{c \in V | (c, d) \in E\}. \quad (9)$$

The conditional probability of Y_c decides the joint distribution of Y as

$$P(Y_c | Y_v - Y_c) = P(Y_c | Y_{N_c}). \quad (10)$$

To prosper the construction, the graph G absorbs a Gibbs distribution all over the maximum cliques (C) in G :

$$P(y) = \prod_{m \in C} \psi_m(y_m) = \frac{1}{Z} e^{-1/T \sum_{m \in C} V_m(y_m)}, \quad (11)$$

where Z is the partition function. Therefore, equation (11) can be rewritten as

$$P(y) = \frac{1}{Z} e^{-1/TU(y)}, \quad (12)$$

where T is the temperature, whose value is generally 1, and $U(y) = \sum_{m \in C} V_m(y_m)$ represents the energy.

Markov models depict the stochastic method that is represented by a graph made of circles has an acute advantage of not considering the past states for all upcoming future states for a random alterable dataset such as HSIs. The variants of Markov random fields are adaptive, hierarchical, cascaded, and probabilistic, a blend of Gaussian mixture model, joint sparse representation, transfer learning, etc., whose outcomes are pretty victorious. Hidden Markov random fields are highly suitable for the unsupervised classification of HSIs where the model parameters are estimated to make each pixel belong to its appropriate cluster [58], leading to the precise classification. Table 3 lists out the research carried out so far for the classification purpose of HSI employing MRF.

4.4. Extreme Learning Machine (ELM). An efficacious learning algorithm based on single hidden layer feedforward neural network (SLFNN), it is applied to classify patterns and regression. Let $(x_i, p_i) \in \mathbf{R}^n \times \mathbf{R}^m$ be N arbitrarily perceptible samples where $x_i = [x_{i1}, \dots, x_{in}]^T \in \mathbf{R}^n$ and $p_i = [p_{i1}, \dots, p_{im}]^T \in \mathbf{R}^m$ [72]. The standard SLFNN having \hat{N} hidden nodes and $f(x)$ as activation function is approached mathematically as

$$\sum_{i=1}^{\hat{N}} \alpha_i f_i(x_i) = \sum_{i=1}^{\hat{N}} \alpha_i f(w_i \cdot x_j + b_i) = O_j; \quad j = 1, \dots, N. \quad (13)$$

TABLE 2: Summary of review of HSI classification using sparse representation.

Year	Method used	Dataset and COA	Research remarks and future scope
2013	Kernel sparse representation classification (KSRC) [45]	IP—96.8%, UP—98.34%, KSC—98.95%	Lacks in devising automatic window size collection of spatial image quality, and filtering degree of class spatial relations
2014	Multiscale adaptive sparse representation (MASR) [46]	UP—98.47%, IP—98.43%, SV—97.33%	MASR outperformed the JSRM single-scale approach and several other classifiers on classification maps and accuracy The structural dictionary desired to be more inclusive and trained by discriminative learning algorithms
2015	Sparse multinomial logistic regression (SMLR) [47]	IP—97.71%, UP—98.69%	Being a pixelwise supervised method, its performance is better than other contemporary methods The model can be improved via more technical validations, exploitation of MRF, and structured sparsity-inducing norm that enhances the interpretability, stability, and identity of the model learned
2015	Super-pixel-based discriminative sparse model (SBDSM) [377]	IP—97.12%, SV—99.37%, UP—97.33%, Washington DC mall—96.84%	The advantages of this model lie in harnessing spatial contexts effectively through the super-pixel concept, which is better in performance speed and classification accuracy Determination of a supplementary and systematic way to adjust the count of super-pixels to various conditions and apply SR to other remote sensing practices
2015	Shape-adaptive joint sparse representation classification (SAJSRC) [48]	IP—98.45%, UP—98.16%, SV—98.53%	Local area shape-adapted for every test pixel rather than a fixed square window for adaptive exploration of spatial PCs, making the method outperforms other corresponding methods Region searching based on shape-adaption can be used instead of the reduced dimensional map to reconnoiter complete spatial information of the actual HSI
2017	Multiple-feature-based adaptive sparse representation (MFASR) [49]	IP—97.99%, UP—98.39%, Washington DC mall—97.26%	SA regions' full utilization of all embedded joint features makes the method superior to some cutting-edge approaches Enhancement of the proposed method in the future by selecting features automatically and improving dictionary learning to reduce the computational cost
2018	Weighted joint nearest neighbor and joint sparse representation (WJNN-JSR) [50]	UP—97.42%, IP—93.95%, SV—95.61%, Pavia center—99.27%	The model was improved using the Gaussian weighted method and incorporates the conventional test pixel area to achieve a new measure of classification knowledge: The Euclidean-weighted joint size Creating more effective approaches to applying the system and further increasing classification accuracy are taken as future work
2019	Log-Euclidean kernel-based joint sparse representation (LogEKJSR) [51]	IP—97.25%, UP—99.06%, SV—99.36%	Specializes in extracting covariance traits from a spatial square neighborhood to calculate the analogy of matrices with covariances employing the conventional Gaussian form of Kernel Creation of adaptive local regions using super-pixel segmentation methods and learning the required kernel using multiple kernel learning methods

TABLE 2: Continued.

Year	Method used	Dataset and COA	Research remarks and future scope
2019	Multiscale super-pixels and guided filter (MSS-GF) [52]	IP—97.58%, UP—99.17%	Effective spatial and edge details in his, various regional scales to build MSSs to acquire accurate spatial information, and GF improved the classification maps for near-edge misclassifications Additional applications of efficient methods to extract local features and segment super-pixels are added as future work
2019	Joint sparse representation—self-paced learning (JSR-SPL) [53]	IP—96.60%, SV—98.98%	The findings are more precise and reliable than other JSR methods
2019	Maximum-likelihood estimation based JSR (MLEJSR) [54]	IP—96.69%, SV—98.91%, KSC—97.13%	The model is reliable in terms of outliers
2020	Global spatial and local spectral similarity-based manifold learning-group sparse representation-based classifier (GSLs-ML-GSRC) [55]	UP—93.42%, Washington DC mall—91.64%, SV—93.79%	The said fusion makes the method outperform other contemporary methods focused on nonlocal or local similarities
2020	Sparse-adaptive hypergraph discriminant analysis (SAHDA) [56]	Washington DC mall—95.28%	Effectively depict the multiple complicated aspects of the HSI and will be considered for future spatial knowledge

Here, $w_i = [w_{i1}, \dots, w_{im}]^T$ gives the weight vector establishing the connection between input nodes and i^{th} is the hidden node and $\alpha_i = [\alpha_{i1}, \dots, \alpha_{im}]^T$ represents the weight vector connecting between output node O_j with the i^{th} hidden node, and $w_i \cdot x_j$ represents the inner product. The zero error for N samples can be written in the matrix form as

$A\alpha = P$, where $A(w_1, \dots, w_{\hat{N}}, b_1, \dots, b_{\hat{N}}, x_1, \dots, x_N)$ is the neural network hidden layer output matrix, and the i^{th} is hidden node output with respect to x_1, \dots, x_N ; the i^{th} column of A represents x_N inputs. The training of SLFNN is based on finding specific α , w_i , and b_i ($i = 1, \dots, \hat{N}$) [73] such that

$$\|A(w_1, \dots, w_{\hat{N}}, b_1, \dots, b_{\hat{N}}, x_1, \dots, x_N)\alpha - P\| = \min_{w, \alpha, b} \|A(w_1, \dots, w_{\hat{N}}, b_1, \dots, b_{\hat{N}}, x_1, \dots, x_N)\alpha - P\|. \quad (14)$$

This equation denotes the cost function with a depreciation. By using gradient-based algorithms, the set of weights (α_i, w_i) and biases b_i are attuned with epochs as

$$w_k = w_{k-1} - \eta \frac{\delta U(W)}{\delta W}; \quad (15)$$

$$U = \sum_{k=1}^N \left(\sum_{j=1}^{\hat{N}} \alpha_j f(w_j \cdot x_k + b_j) - P_k \right)^2.$$

The learning rate η must be accurate for better convergence and $\hat{N} \ll N$ for better generalization performance.

Extreme learning methods proposed overcoming the disadvantage of a single hidden layer feedforward neural network and improving learning ability and generalization performance. It is a supervised method but is highly recommended to get an extension to its semi-supervised and unsupervised versions for dealing with the huge amount of data such as HSIs, which are primarily unlabeled and suffering from lack of training samples. Great potential lies with its other variants than those mentioned here, [74] of ELM, like two-hidden layer ELM, multilayer ELM, feature mapping-based ELM, incremental ELM, and deep ELM to become superior and achieve victorious precision in classifying

HSIs. Table 4 underneath provides the summary of the research executed so far for the classification purpose of HSI utilizing ELM.

4.5. Active Learning (AL). It is a special type of the supervised ML approach to build a high-performance classifier while minimizing the size of the training dataset by actively selecting valuable data points. The general structure of AL can be understood from Figure 6. There are three categories of AL—stream-based selective sampling, that is, where each unlabeled dataset is enquired for a certain label whether to assign a query or not; pool-based sampling; that is, the whole dataset is under consideration before selecting the best set of queries; and membership query synthesis; that is, it involves data augmentation to create user selected labeling. The decision to select the most informative data points depends on the uncertainty measure used in the selection. In an active learning scenario, the most informative data points are those the classifier is least sure about. The uncertainty measures for datapoints x [88] are

Least Confidence (LC): responsible for selecting the classifier's data point is least certain about the chosen class. With y^* as the most likely label sequence and ϕ as the learning model, LC is represented as

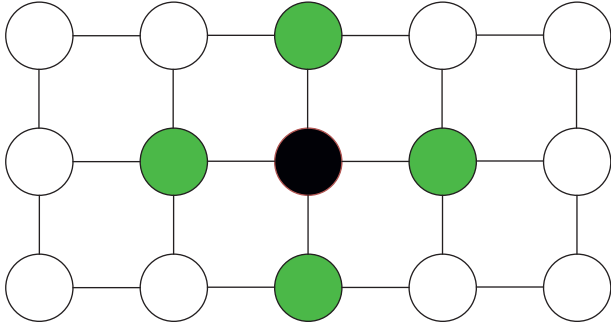


FIGURE 5: Given the green nodes, the black node is independent of other nodes.

$$S_{LC}(x) = 1 - P(y^* | x, \phi). \quad (16)$$

Smallest Margin Uncertainty (SMU): Represents the difference between classification probability of the most likely class (y_1^*) and that of the second-best class (y_2^*), written mathematically as:

$$S_{SMU}(x) = P_{\phi}(y_1^* | x) - P_{\phi}(y_2^* | x). \quad (17)$$

Largest Margin Uncertainty (LMU): Represents the difference between classification probability of most likely class (y_1^*) and that of the least likely class (y_{\min}), written mathematically as:

$$S_{LMU}(x) = P_{\phi}(y_1^* | x) - P_{\phi}(y_{\min}^* | x). \quad (18)$$

Sequence Entropy (SE): Detects the measure of disorder in a system; higher the entropy implies a more disordered condition. The denotation of SE is

$$S_{SE}(x) = - \sum_{\hat{y}} P(\hat{y} | x; \phi) \log P(\hat{y} | x; \phi), \quad (19)$$

with \hat{y} ranging over all possible label sequences for input x .

Although not considered customary and coherent, AL is pretty much capable of reducing human effort, time, and processing cost for a large batch of unlabeled data. This method relies on prioritizing data that needs to be labeled in a huge pool of unlabeled data to have the highest impact on training. A desired supervised model keeps on being trained through active queries and improvising itself to predict the class for each remaining data point. AL is advantageous for its dynamic and incremental approach to training the model so that it learns the most suitable label for each data cluster [89]. Table 5 lists out the research performed so far for the classification purpose of HSI using AL.

4.6. Deep Learning (DL). Deep learning is the most renowned ML technology in application and accuracy terms. Although it is considered the next tread of ML, it also lends concepts from artificial intelligence. DL is the mother of algorithms that resemble human brain simulations, that is, creativity, enhanced analysis, and proper decision-making, based on pure or hybrid large networks for any given real-life

problem. It has enhanced the throughput of computer-based, especially unsupervised snags for the practical technology-based applications such as automated translation of machines, image reconstructions and classifications, computer vision, and automated analysis. [104] The basic structure of any DL model possesses a three-type-layered architecture: it contains one input layer through which input data are fed to the next layer(s) known as the intermediate hidden layer responsible for all the computations based on the problem given, which passes its generated data to the final layer, that is, the output layer, which provides the desired ultimate output. The steps involved in DL models are as follows: having proper knowledge and understanding of the problem, collecting the input database, selecting the most appropriate algorithm, training the model with the sample source database, and finally testing the target database [105].

DL models are more efficient and advantageous over other ML models due to the following reasons [19]:

- (1) The capability to extract hidden and complicated structures from raw data is inextricably linked to their ability to represent the internal representation and generalize any form of knowledge.
- (2) They have a wide range of data types that they can accommodate, for example, 2D imagery data and complex 3D data such as medical imagery and remote sensing. In addition, they can use HSI data's spectral and spatial domains in both standalone and linked ways [106–108].
- (3) They provide architects a lot of versatility in terms of layer types, blocks, units, and depth.
- (4) Furthermore, its learning approach can be tailored to various learning strategies, from unsupervised to supervised, with intermediate strategy.
- (5) Additionally, developments in processing techniques, including batch partitioning and high-performance computation, especially on distributed and parallel architecture, have enabled DL models to find better opportunities and solutions when coping with enormous volumes of data [109].

The models that are broadly used for HSI classification are described as follows.

- (a) **Autoencoder (AE)**: AEs are the fundamental unsupervised deep model based on the back-propagation rule. AEs consist of two fragments: encoder, connecting the input vector to the hidden layer by a weight matrix; decoder, formed by the hidden layer output via a reconstruction vector tied by a specific weight matrix. SAEs are AEs with multiple hidden layers where the production of every hidden layer is fed to the successive hidden layer as input. It comprises three steps: (1) first AE trained to fetch the learned feature vector; (2) the former layer's feature vector is taken as input to the next layer, and this process is redone till the completion of training;

TABLE 3: Summary of review of HSI classification using MRF.

Year	Method used	Dataset and COA	Research remarks and future scope
2011	Adaptive-MRF (a-MRF) [59]	IP—92.55%	Handles homogeneous problem of “salt and pepper” areas and the possibility of overcorrection impact on class boundaries
2014	Hidden MRF and SVM (HMRF-SVM) [60]	IP—90.50%, SV—97.24%	Outperforms SVM and improves overall accuracy outcomes by nearly 8% and 3.2%, respectively
2014	Probabilistic SR with MRF-based multiple linear logistic (PSR-MLL) [61]	IP—97.8%, UP—99.1%, Pavia center—99.4%	Exceeds other modern contemporary methods in terms of accuracy
2014	MRF with Gaussian mixture model (GMM-MRF) [62]	UP(LFDA-GMM-MRF)—90.88% UP(LPNMF-GMM-MRF)—94.96%	Advantageous for a vast range of operating conditions and spatial-spectral information to preserve multimodal statistics GMM classificatory distributions are to be considered in the future
2011	MRF with sparse multinomial logistic regression classifier—spatially adaptive total variation regularization (MRF-SMLR-SpATV) [63]	UP—90.01%, IP—97.85%, Pavia center—99.23%	Efficient time complexity of the model Improvisation of the model by implementing GPU and learning dictionaries are the future agendas
2016	Multitask joint sparse representation (MJSR) and a stepwise Markov random field framework (MSMRF) [64]	IP—92.11%, UP—92.52%	The gradual optimization explores the spatial correlation, which significantly improves the effectivity and accuracy of the classification
2016	MRF with hierarchical statistical region merging (HSRM) [65]	SVMMRF-HSRM: IP—93.10%, SV—99.15%, UP— 86.52%; MLRsubMRF-HSRM-IP—82.60%, SV—88.16%, UP—95.52%	Better solution to the technique of majority voting that suffers from the problem of scale choice Considering the spatial features in the spatial prior model of objects of the different groups in the future
2018	Integration of optimum dictionary learning with extended hidden Markov random field (ODL-EMHRF) [66]	ODL-EMHRF-ML-IP—98.56%, UP—99.63%; ODL-EMHRF-EM- IP—98.47%, UP—99.58%	The method has been proven to be better than SVM-associated EMRF
2018	Label-dependent spectral mixture model (LSMM) fused with MRF (LSMM-MRF) [67]	The Konka image—94.19%, the shipping scene—66.45%	Efficient unsupervised classification strategy that considers spectral information in mixed pixels and the impact of spatial correlation Enhanced theoretical derivations of EM steps
2019	Adaptive interclass-pair penalty and spectral similarity information (aICP2-SSI) along with MRF and SVM [68]	UP—98.10%, SV—96.40%, IP— 96.14%	Outperforms other MRF-based methods More efficient edge-preserving strategies, more spectral similitude, and class separable calculation methods as future research
2019	Cascaded version of MRF (CMRF) [69]	IP—98.56%, Botswana—99.32%, KSC—99.24%	Backpropagation tunes the model parameters and least computation expenses
2020	Fusion of transfer learning and MRF (TL-MRF) [70]	IP—93.89%, UP—91.79%	TL is taken to be very effective for HSI classification Future research for reducing the number of calculations involved in the existing
2020	MRF with capsule net (caps-MRF) [71]	IP—98.52%, SV—99.74%, Pavia center—99.84%	Ensures that relevant information is preserved, and the spatial constraint of the MRF helps achieve more precise model convergence The combination of CapsNet with several postclassification techniques

TABLE 4: Summary of review of HSI classification using ELM.

Year	Method used	Dataset and COA	Research remarks and future scope
2014	Ensemble extreme learning machines (E ² LM)-bagging-based ELMs (BagELMs) and AdaBoost-based (BoostELMs) [72]	UP—94.3%, KSC—97.71%, SV—97.19%	BoostELM performs better than kernel and other EL methods Performance of other differential or nondifferentiable activation functions
2015	Kernel-based ELM—composite kernel (KELM-CK) [75]	IP—95.9%, UP—93.5%, SV—96.4%	Outperforms other SVM-CK-based models
2015	ELM's two-level fusions: feature-level fusion (FF-ELM) and mixing ELM classifier two levels of fusions: feature-level fusion (FF-ELM) [76]	FF-ELM: UP—98.11%, IP—92.93%, SV—99.12%; DF-ELM—UP—99.25%, IP—93.58%, SV—99.63%	Outperforms basic ELM models
2016	Hierarchical local-receptive-field-based ELM (HL-ELM) [77]	IP—98.36%, UP—98.59%	Surpasses other ELM methods in terms of accuracy and training speed
2017	Genetic-firefly algorithm with ELM (3FA-ELM) [78]	HyDice DC mall—97.36%, HyMap—95.58%	Low complexity (ELM), better adaptability, and searching capability (FA) Execution time needs to be reduced in future
2017	Local receptive fields-based kernel ELM (LRF-KELM) [79]	IP—98.29%	Outperforms other ELM models
2017	Distributed KELM based on MapReduce framework with Gabor filtering (DK-Gabor-ELMM) [80]	IP—92.8%, UP—98.8%	Outperforms other ELM models
2017	Loopy belief propagation with ELM (ELM-LBP) [81]	IP—97.29%	Efficient time complexity
2018	Mean filtering with RBF-based KELM (MF-KELM) [82]	IP—98.52%	The model offers the most negligible computational hazard
2018	Augmented sparse multinomial logistic ELM (ASMLELM) [83]	IP—98.85%, UP—99.71%, SV—98.92%	Improved classification accuracy by extended multi-attribute profiles and more SR
2018	ELM with enhanced composite feature (ELM-ECF) [84]	IP—98.8%, UP—99.7%, SV—99.5%	Low complexity and multiscale spatial feature for better accuracy Incorporate feature-fusion technology
2019	Local block multilayer sparse ELM (LBMS-ELM) [85]	IP—89.31%, UP—89.47%, SV—90.03%	Performs anomaly and target detection. Reduced computational overhead and increased classification accuracy by inverse free; saliency detection and gravitational search
2019	ELM-based heterogeneous domain adaptation (EHDA) [25]	HU-DC —97.51%, UP-DC —96.63%, UP-HU —97.53%	Outperforms other HDA methods. Invariant feature selection
2019	Spectral-spatial domain-specific convolutional deep ELM (S ² CDELm) [86]	IP—97.42%, UP—99.72%	Easy construction with high training-testing speed Merge of DL with ELM
2020	Cumulative variation weights and comprehensive evaluated ELM (CVW-CEELM) [87]	IP—98.5%, UP—99.4%	Accuracy achieved due to the weight determination of multiple weak classifiers. Multiscale neighborhood choice and optimized feature selection

(3) backpropagation is used after all the hidden layers have been trained to reduce the cost function and to update the weights is done with a named training set to obtain fine-tuning [110]. The architecture of SAE is depicted in Figure 7.

Let $x_n \in \mathbf{R}^m$; $n = 1, 2, \dots, N$ represent the unlabeled input dataset, E_n be the hidden encoder vector computed by x_n , and y_n be the decoder vector of the output layer [111].

$$\text{Encoder} : E_n = g(W_i x_n + b_i); \quad (20)$$

g -> encoding function, W_i -> encoder weight matrix, b_i -> encoder bias vector.

$$\text{Decoder} : y_n = f(W_j E_n + b_j); \quad (21)$$

f -> decoding function, W_j -> decoder weight matrix, b_j -> decoder bias vector.

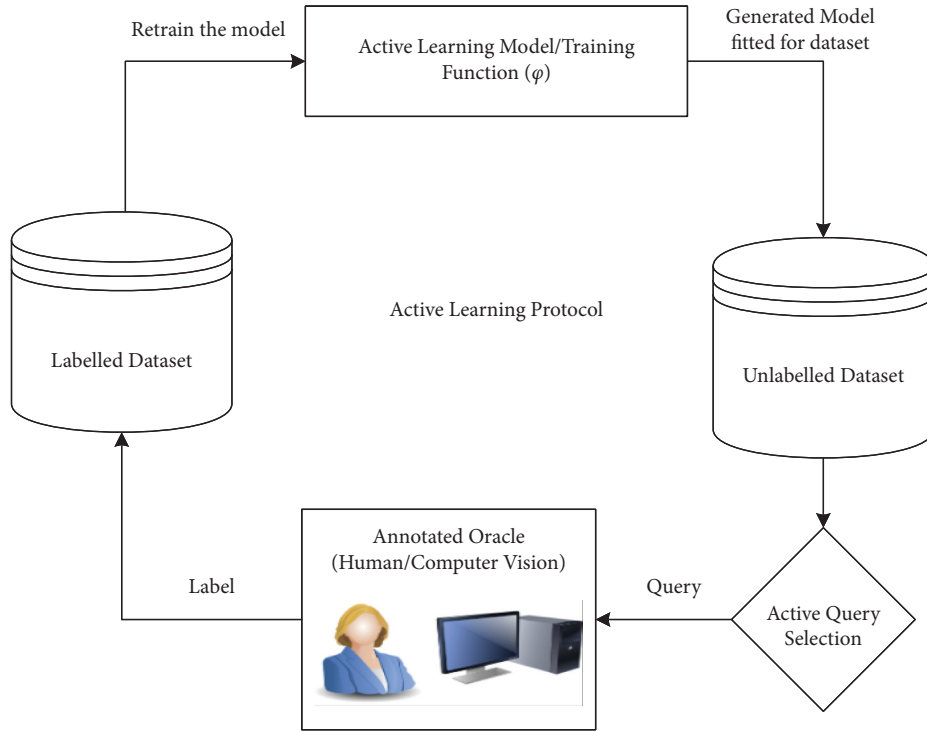


FIGURE 6: Principle of active learning.

The reconstruction error in SAE is denoted as

$$\Phi(\Theta) = \arg \min_{\theta, \theta'} \frac{1}{N} \sum_{k=1}^N L(x^k, y^k), \text{ where the Loss function is } L(x^k, y^k) = \|x - y\|^2. \quad (22)$$

AEs are unsupervised neural networks that embed several convolutional hidden layers based on nonlinear activation functions and transformations [112]. There are high risks of data loss during training, but it handles the model well for specific data types through specialized training. There are AEs for every purpose such as convolutional, sparse, variational, deep, contractive, and denoising applied for data compression, noise removal, feature extraction, image augmenting, and image coloring. AE inevitably provides a vast platform for further research on its various applicability and its capability to participate in hybridization. Table 6 describes a few research works in the aspect of AEs.

- (b) *Convolutional Neural Network (CNN)*: It is a famous deep neural network that works like a human visual cortex with many interconnected layers applied widely in image, speech, and signal processing. It assigns learnable and modifiable weights and biases to the input image to identify various objects or patterns with differentiable features. As shown in

Figure 8, each layer of CNN possesses filtering capabilities with ascending complexities: the first layer learns filtering corners and edges; intermediate layers learn object parts filtering; and the last layer learns filtering out the entire object in different locations and shapes. The comparison between the layers in terms of several parameters is shown in Table 7. It consists of four layers [117, 118]:

- (1) *Convolution*: This operation is the cause of the naming of CNN, that is, a dot product of the original pixel values with weights identified in the filter or kernel of the image. The findings are compiled into one number representing all the pixels found in the filter. Assuming \mathbf{I} be the hyper-input-cube of dimension $p \times q \times r$ where $p \times q$ denotes the spatial size of \mathbf{I} with r number of bands, and i_k is the k th feature map of \mathbf{I} . Let d number of filters be present in each convolutional layer, and weight W_m and bias b_m represent the m th filter. The m th convolutional layer output with transformation function g is denoted as

TABLE 5: Summary of review of HSI classification using active learning.

Year	Method used	Dataset and COA	Research remarks and future scope
2008	AL with expectation-maximization-binary hierarchical classifier (BHC-EM-AL) and maximum-likelihood (ML-EM-AL) [90]	Range: KSC-90-96%, Botswana—94-98%	Better learning levels than the random choice of data points and an entropy-based AL Measurement of the efficacy of the active learning-based knowledge transfer approach while systematically increasing the spatial/temporal segregation of the data sources
2010	Semi-supervised-segmentation with AL and multinomial logistic regression (MLR-AL) [91]	IP—79.90%, SV—97.47%	Innovative mechanisms for selecting unlabeled training samples automatically, AL to enhance segmentation results Testing the segmentation in various scenarios influenced by limited a priori accessibility of training images
2013	Maximizer of the posterior marginal by loopy belief propagation with AL (MPM-LBP-AL) [92]	IP—94.76%, UP—85.78%	Improved accuracy than previous AL applications Use parallel-computer-architectures such as commodity—clusters or GPUs to build computationally proficient implementation
2015	Hybrid AL-MRF, that is, uncertainly sampling breaking ties (MRF-AL-BT), passive selection approach random sampling (MRF-AL-RS), and the combination (MRF-AL-BT + RS) [93]	IP—94.76%, UP—85.78% (MRF-AL-RS provides the highest accuracies)	Outperforms conventional AL and SVM AL methods due to MRF regularization and pixelwise output Merge the model with other effective AL methods and test them with a limited number of training samples
2015	Integration of AL and Gaussian process classifier (GP-AL) [94]	IP—89.49%, Pavia center—98.22%	Empirical automation of AL achieves reasonable accuracy Adding diversity criterion to the heuristics and contextual information with the model and reducing computation time
2016	AL with hierarchical segmentation (HSeg) tree: adding features and adding samples (Adseg_AddFeat + AddSamp) [95]	IP—82.77%, UP—92.23%	Outruns several baseline methods-selecting appropriate training data from already existing labeled datasets and potentially decreasing manual laboratory labeling Reduce the computational time that limits its applicability on large-scale datasets
2016	Multiview 3D redundant discrete wavelet transform-based AL (3D-RDWT-MV-AL) [96]	HU—99%, KSC—99.8%, UP—95%, IP—90%	The precious method as a combination of an initial process with AL, improved classification
2017	Discovering representativeness and discriminativeness by semi-supervised active learning (DRDbSSAL) [97]	Botswana—97.03%, KSC—93.47%, UP—93.03%, IP—88.03%	Novel approach with efficient accuracy
2017	Multicriteria AL [98]	KSC—99.71%, UP—99.66%, IP—99.44%	Surpasses other existing AL methods regarding stability, accuracy, robustness, and computational hazard A multi-objective optimization strategy and the usage of advanced attribute-based profile features
2018	Feature-driven AL associated with morphological profiles and Gabor filter [99]	IP—99.5%, UP—99.84%, KSC—99.53% (Gabor-BT)	A discriminative feature space is designed to gather helpful information into restricted samples
2018	Multiview intensity-based AL (MVAL)-multiview intensity-based query-representative strategy (MVIQ-R) [100]	UP—98%, Botswana—99.5%, KSC—99.9%, IP—95%	Focus on pixel intensity obtains unique feature and hence better performance Selection of combination of optimal attribute features
2019	Super-pixel with density peak augmentation (DPA)-based semi-supervised AL (SDP-SSAL) [101]	IP—90.08%, UP—85.61%	Novel approach proposed based on super-pixels density metric Development of a pixelwise solution to produce super-pixel-based neighborhoods

TABLE 5: Continued.

Year	Method used	Dataset and COA	Research remarks and future scope
2020	Adaptive multiview ensemble spectral classifier and hierarchical segmentation (Ad-MVEnC_Spec + Hseg) [102]	KSC—97.63%, IP—87.1%, HU—93.3%	Enhancement in the view sufficiency, and promotion of the disagreement level by the dynamic view, provides lower computational complexity due to parallel computing
2020	Spectral-spatial feature fusion using spatial coordinates-based AL (SSFFSC-AL) [103]	IP—100%, UP—98.43%	High running speed can successfully address the “salt and pepper” phenomenon but drops a few if similar class samples are distributed in different regions differently The sampling weight parameter conversion to an adaptive parameter is adjusted adaptively as the training samples are modified

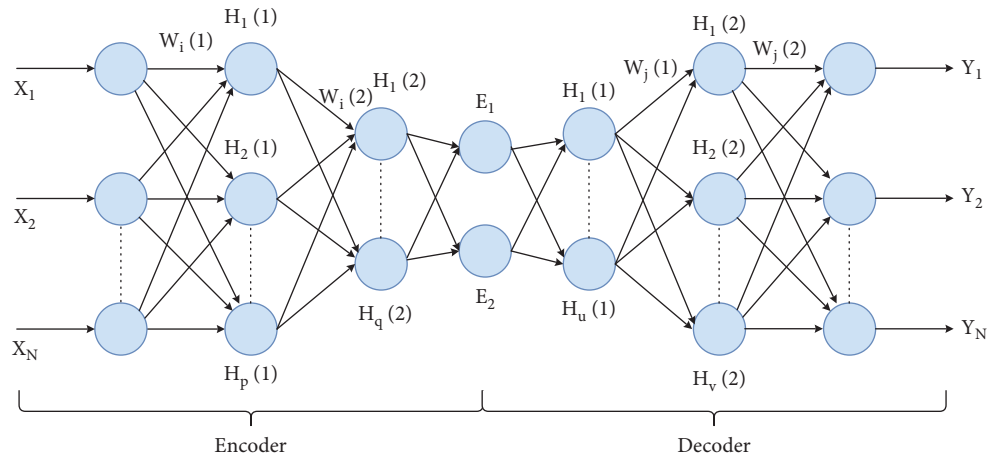


FIGURE 7: The network structure of stacked autoencoders; input X-to-E is the encoding phase; E-to-output Y is the decoding phase.

TABLE 6: Summary of the review of HSI classification using deep learning—AE.

Year	Method used	Dataset and COA	Research remarks and future scope
2013	Autoencoders (AE) [110]	Error rate: KSC—4%, Pavia city—14.36%	This article opened a considerable doorway of research, including other deep models for better accuracy
2014	Stacked autoencoder and logistic regression (SAE-LR) [113]	KSC—98.76%, Pavia city—98.52%	Highly accurate in comparison to RBF-SVM and performs testing in optimized time limit than SVM or KNN but fails in training time efficiency
2016	Spatial updated deep AE with collaborative representation-based classifier (SDAE-CR) [114]	IP—99.22%, Pavia center—99.9%, Botswana—99.88%	Highly structured in extracting high specialty deep features and not the hand-crafted ones and accurate Improving the deep network architecture and selection of parameters
2019	Compact and discriminative stacked autoencoder (CDSAE) [115]	UP—97.59%, IP—95.81%, SV—96.07%	Efficient in dealing with feature space in low dimension, but the computation cost is high as per architecture size
2021	Stacked autoencoder with distance-based spatial-spectral vector [116]	SV—97.93%, UP—99.34%, surrey—94.31%	Augmentation of EMAP features with the geometrically allocated spatial-spectral feature vectors achieves excellent results. Better tuning of hyperparameter and more powerful computational tool required Improving the training model to become unified and classified in a more generalized and accurate way

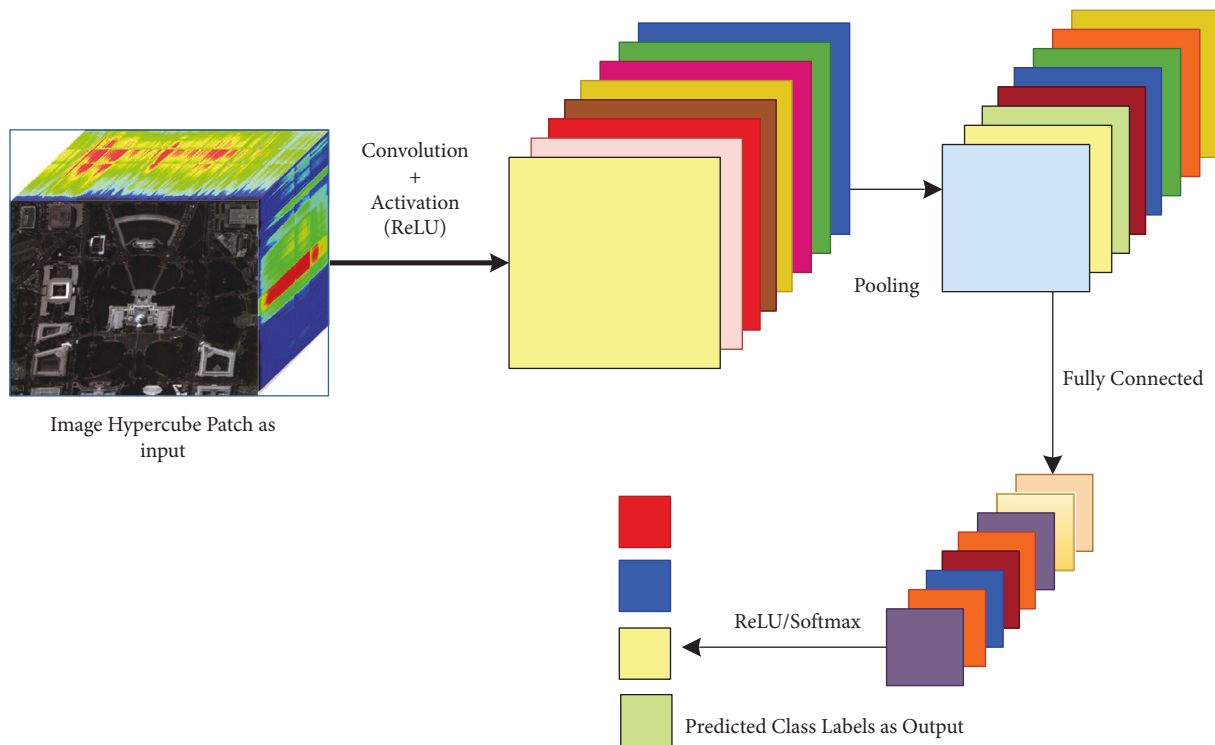


FIGURE 8: The CNN architecture deploying the layers.

TABLE 7: Comparison of convolutional layers.

Arguments	Convolution layer	Pooling layer	Fully connected layer
Input	(i) 3D-cube, preceding set of feature maps	(i) 3D-cube, preceding set of feature maps	(i) Flattened-3d-cube, preceding set of feature maps
	(i) Kernel counts	(i) Stride	(i) Number of nodes
	(ii) Kernel size	(ii) Size of window	(ii) Activation function: selected based on the role of the layer. For aggregating info-ReLU. For producing final classification—softmax
Parameters	(iii) Activation function (ReLU) (iv) Stride (v) Padding (vi) Type and value of regularization		
	(i) Application of filters made of small kernels to extricate features	(i) Reduction of dimensionality	(i) Aggregate information from final feature maps
	(ii) Learning	(ii) Extraction of the maximum of a region average	(ii) Generate final classification
Action	(iii) One bias per filter (iv) Application of activation function on each feature map value	(iii) Sliding window framework	
Output	(i) 3D-cube, a 2D-map per filter	(i) 3D-cube, a 2D-map per filter, reduced spatial dimensions	(i) 3D-cube, a 2D-map per filter

TABLE 8: Summary of review of HSI classification using deep learning—CNN.

Year	Method used	Dataset and COA	Research remarks and future scope
2015	Convolutional neural network and multilayer perceptron (CNN-MLP) [120]	Pavia city—99.91%, UP—99.62%, SV—99.53%, IP—98.88%	Far better than SVM, RBF mixed classifiers, the effective convergence rate can be useful for large datasets Detection of human behavior from hyperspectral video sequences
2016	3D-CNN [121]	IP—98.53%, UP—99.66%, KSC—97.07%	A landmark in terms of quality and overall performance Mapping performance to be accelerated by postclassification processing
2016	Spectral-spatial feature-based classification (SSFC) [122]	Pavia center—99.87%, UP—96.98%	Highly accurate than other methods Inclusion of optimal observation scale for improved outcome
2016	CNN-based simple linear iterative clustering (SLIC-CNN) [123]	KSC—100%, UP—99.64, IP—97.24%	Deals with a limited dataset use spectral and local-spatial probabilities as an enhanced estimate in the Bayesian inference
2017	Pixel-pair feature enhanced deep CNN (CNN-PPF) [124]	IP—94.34%, SV—94.8%, UP—96.48%	Overcomes the significant parameter and bulk-data problems of DL, PPFs make the system unique and reliable, and voting strategy makes the more enhanced evaluations in classification
2017	Multiscale 3D deep convolutional neural network (M3D-DCNN) [125]	IP—97.61%, UP—98.49%, SV—97.24%	Outperforms popular methods like RBF-SVM and combinations of CNNs Removing data limitations and improving the network architecture
2018	2D-CNN, 3D-CNN, recurrent 2D-CNN (R-2D-CNN), and recurrent 3D-CNN (R-3D-CNN) [126]	IP-99.5%, UP—99.97%, Botswana—99.38%, PaviaC—96.79%, SV—99.8%, KSC—99.85%	R-3D-CNN outperforms all other CNNs mentioned and proves to be very potent in both fast convergence and feature extraction but suffers from the limited sample problem Applying prior knowledge and transfer learning
2019	3D lightweight convolutional neural network (CNN) (3D-LWNet) [127]	UP—99.4%, IP—98.87%, KSC—98.22%	Provides irrelevance to the sources of data Architecture is to be improvised by intelligent algorithms
2020	Hybrid spectral CNN (HybridSN) [128]	IP—99.75%, UP—99.98%, SV—100%	Removes the shortfalls of passing over the essential spectral bands and complex, the tedious structure of 2D-CNN and 3D-CNN exclusively and outruns all other contemporary CNN methods superiorly, like SSRN and M-3D-CNN
2020	Heterogeneous TL based on CNN with attention mechanism (HT-CNN-attention) [129]	SV—99%, UP—97.78%, KSC—99.56%, IP—96.99%	Efficient approach regardless of the sample selection strategies chosen
2020	Quantum genetic-optimized SR based CNN (QGASR-CNN) [27]	UP—91.6%, IP—94.1%	With enhanced accuracy, overfitting and “salt-and-pepper” noise are resolved Improvement of operational performance by the relation between feature mapping and selection of parameters
2020	Rotation-equivariant CNN2D (reCNN2D) [130]	IP—97.78%, UP—98.89, SV—98.18%	Provides robustness and optimal generalization and accuracy without any data augmentation
2020	Spectral-spatial dense connectivity-attention 3D-CNN (SSDANet) [131]	UP—99.97%, IP— 99.29%	Higher accuracy but high computational hazard Optimization by using other efficient algorithms

$$Y_m = \sum_{k=1}^r g(i_k \cdot W_m + b_m); m = 1, 2, \dots, d. \quad (23)$$

(2) *Activation*: The convolution layer produces a matrix significantly smaller than the actual image. The matrix is passed through an activation layer (generally rectified linear unit, aka ReLU),

adding nonlinearity that enables the network to train itself through backpropagation.

(3) *Pooling*: It is the method of even more down-sampling and reduction of the matrix size. A filter is applied over the results obtained by the previous layer and chooses a number from each set of values (generally the maximum, the max-

pooling), which allows the network to train much more quickly, concentrating on the most valuable information in each image feature. For an $m \times m$ square window neighbor S with N elements and z_{ij} activation value concerning (i, j) location, the average pooling is formulated as

$$T = \frac{1}{N} \sum_{(i,j) \in S} z_{ij}. \quad (24)$$

- (4) *Fully Connected (FC)*: A typical perceptron structure with multilayers. The input is a single-dimensional vector representing the output of the layers above. Its output is a probability list for the various possible labels attached to the image. Classification decision is the mark that receives the highest likelihood. It is mathematically represented with transformation function g , for N samples of inputs with X'' and Y'' being the outputs having W being the weight matrix and b , the bias constant, is as follows:

$$Y'' = \sum_{j=1}^N g(WX'' + b). \quad (25)$$

CNN is the most method-in-demand and widely explored model among all DL models. The functional unit of convolutional layers is kernels that expertise in extricating the most relevant and enriched spatial and spectral features from the given dataset through automated filtering by convolution operation [119]. It provides an intense description of the whereabouts of CNNs. The most popular ones are attention-based CNN, ResNet, CapsNet, LeNet, AlexNet, VGG, etc. Some of them are still unexplored yet in classifying HSI. The detailed research work on CNN for dealing with HSI classification is listed in Table 8.

- (c) *Recurrent Neural Network (RNN)*: DL is a very efficient approach that follows a sequential framework with a definite timestamp t . "Recurrent" refers to performing the same task for each sequence element, with the output depending on the preceding computations. In other words, they have a "memory" that enfolds information about the calculation so far type of neural network, and the output of a particular recurrent neuron is fed backward as input to the same node, which leads the network to efficiently predict the output, represented in Figure 9, where RNN unrolls, that is, show the complete sequence of the entire network structure neuron by neuron. It consists of the following steps:

- (1) $\mathbf{X} = [\dots, x_{t-1}, x_t, x_{t+1}, \dots]$ be the input vector, where x_t represents input at timestamp t .
- (2) h_t is the "memory of the network," the hidden state at timestamp t . Preliminarily, h_{-1} is initialized to zero vector to calculate the first hidden step. h_t being the current step is calculated based

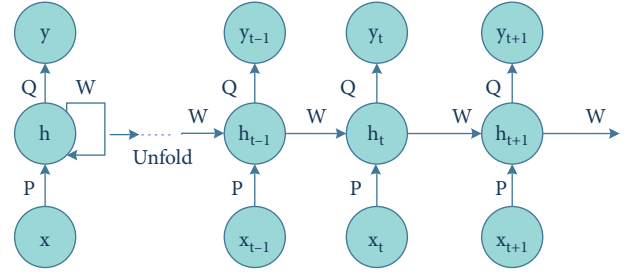


FIGURE 9: The RNN structure with recurrent neurons.

on previously hidden step h_{t-1} , formulated by [132]

$$h_t = f(Px_t + Wh_{t-1}), \quad (26)$$

where f denotes a function of nonlinearity, that is, tanh or ReLU, and W be the weight vector.

- (3) $\mathbf{Y} = [\dots, y_{t-1}, y_t, y_{t+1}, \dots]$ be the output vector, where y_t represents input at timestamp t , generally a softmax function: $y_t = \text{softmax}(Q h_t)$.

RNN is an efficient deep model with large potential. The recurrence looping structure acquainted with RNN enables it to store relevant information about spatial-spectral relationships between the pixels and neighbors. There are several RNN architectures based on inputs/outputs as stated in [133], and based on LSTM, there are five categories [134]. These variates can be well utilized in collaboration with other DL methods such as MRF and PCA to find their accuracy.

The literature studies based on RNN are cataloged in Table 9.

- (d) *Deep Belief Network (DBN)*: DBNs are formed by greedy stacking and training restricted Boltzmann machines (RBMs), an unsupervised learning algorithm based on "contrastive divergence." For neural networks, RBMs suggest taking a probabilistic approach and are thus called stochastic neural networks. Each RBM is made of three parts: a visible unit (input layer), an invisible unit (hidden layer), and a bias unit. The general structure of a DBN is depicted in Figure 10.

For a DBN, the joint distribution of input vector, X with n hidden layers h_n , is defined as [137]

$$P(X, h_1, \dots, h_n) = \left(\prod_{i=0}^{n-2} P(h_i | h_{i+1}) \right) \cdot P(h_{n-1}, h_n), \quad (27)$$

where $X = h_0$, $P(h_{i-1}, h_i)$ is the conditional distribution of the visible units on the hidden RBM units at level i and $P(h_{n-1}, h_n)$ is the hidden-visible joint distribution in top-level RBM. DBN has two phases: the pretraining phase depicts numerous layers of RBM, and fine-tuning phase is simply a feedforward NN.

TABLE 9: Summary of review of HSI classification using deep learning—RNN.

Year	Method used	Dataset and COA	Research remarks and future scope
2017	Gated recurrent unit-based RNN with parametric rectified tanh as activation function (RNN-GRU-pretanh) [132]	UP—88.85%, HU—89.85%, IP—88.63%	An enhanced model that utilizes the intrinsic feature provided by HS pixels with better accuracy than SVM The study is limited to only spectral features Incorporation of deep end-to-end convolutional RNN with both spatial-spectral features
2019	Spectral-spatial cascaded recurrent neural network (SSCasRNN) [135]	IP—91.79%, UP—90.30%	Outruns pure RNN and CNN models due to the perfect placement of convolutional and recurrent layers to explore joint information
2020	Geometry-aware deep RNN (Geo-DRNN) [136]	UP—98.05%, IP—97.77%	Due to encoding the complex geometrical structures, the data lack space Minimization of memory-occupation
2021	2D and 3D spatial attention-driven recurrent feedback convolutional neural network (SARFNN) [28]	IP—99.15%, HU—86.05%	Integrating attention and feedback mechanism with recurrent nets in two layers, 2D and 3D, enables efficient accuracy

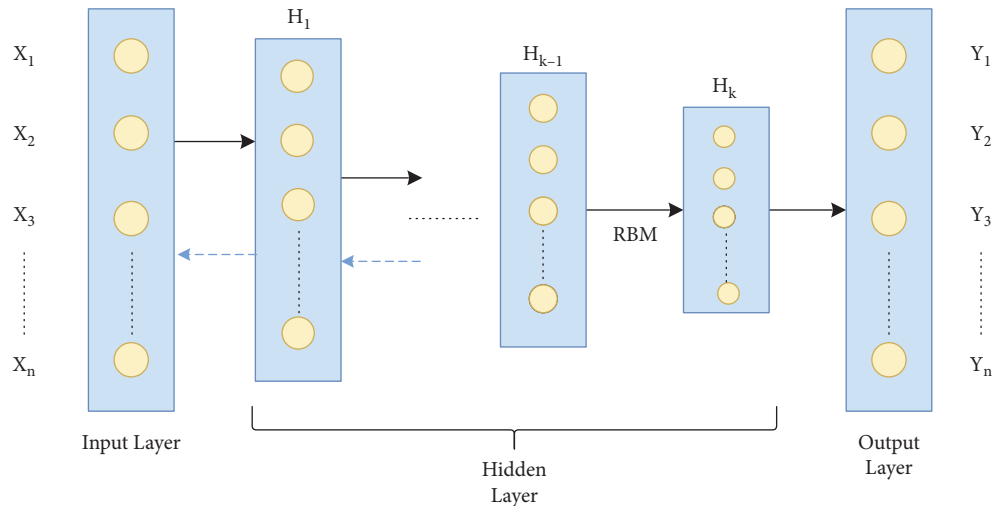


FIGURE 10: The detailed DBN structure.

TABLE 10: Summary of review of HSI classification using deep learning—DBN.

Year	Method used	Dataset and COA	Research remarks
2015	Deep belief network and logistic regression (DBN-LR) [137]	IP—95.95%, Pavia City—99.05%	The drawback in training time complexity, it is super-fast testing, and result generating capability outperforms RBF-SVM with EMP
2019	Spectral-adaptive segmented deep belief network (SAS-DBN) [138]	UP—93.15%, HU—98.35%	Capable of addressing the complexities and other subsidiaries of limited samples
2020	Conjugate gradient update-based DBN (CGDBN) [139]	UP—97.31%	Better approach towards stability and convergence of the training model High time complexity

DBN is the graphical representation that is generative; that is, it creates all distinct outcomes that can be produced for the particular case and learn to disengage a deep hierarchical depiction of the sample training data. DBNs are structurally more capable than RNNs as they lack loops, are pretrained in an unsupervised way, and are computationally eminent for particularly classification problems. Minor modifications or collaborations can improvise DBNs functionally and accuracy. Table 10 depicts a list of works done on DBN.

(e) Generative Adversarial Network (GAN): One of the most recent DL models that are rapidly growing its footsteps in the area of technical research. The GAN model is trained using two kinds of neural networks: the “generative network” or “generator” model that learns to generate new viable samples and the “discriminatory network” or “discriminator,” which learns to discriminate generated instances from existing instances. Discriminative algorithms seek to classify the input data, which is given as a collection of certain features; the algorithm maps feature on

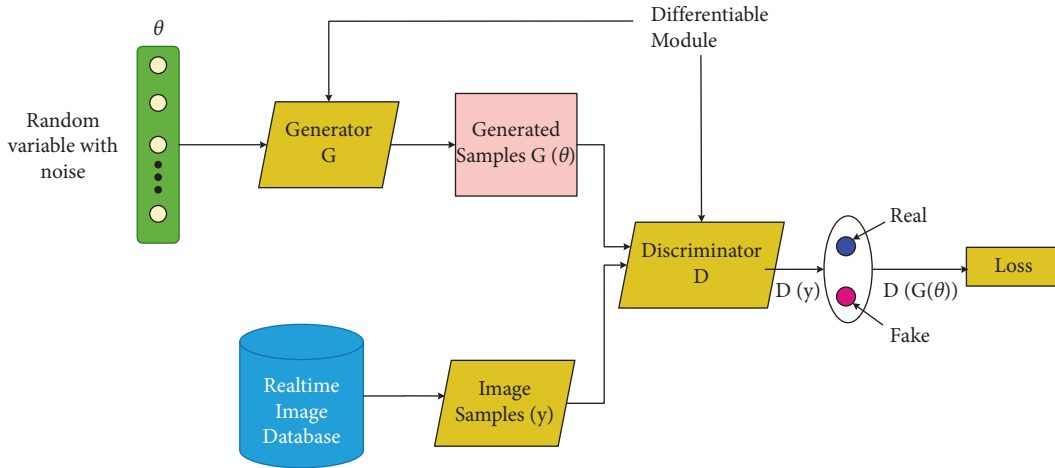


FIGURE 11: The GAN architecture.

labels [140]. In contrast, generative algorithms attempt to construct the input data, which is given with a set of features, and it will not classify it, but it will attempt to create a feature that matches a certain label. The generator tries to get better at deluding the discriminator during the training, and the discriminator tries to grab the counterfeits generated by the generator. Thus, the training procedure is termed adversarial training. The generator and discriminator should be trained against a static opponent, keeping the discriminator constant while training the generator and keeping the generator constant when training the discriminator. That helps to understand the gradients better.

In a GAN model, say D and G denote the discriminator and the generator units that map a noise data space θ to real and original data space x , respectively. $G(\theta)$ denotes the fake output generated by G , and $D(y)$, and $D(G(\theta))$ are D 's output for real and fake training samples, respectively. $P_\theta(\theta)$ and $P_d(y)$ represent the input model distribution and original data distribution, respectively, when $\theta \sim P_\theta$ [141] as shown in Figure 11.

$$\begin{aligned} \text{The loss function for } D: L^{(D)} = & \max[\log(D(y)) \\ & + \log(1 - D(G(\theta)))]. \end{aligned} \quad (28)$$

$$\begin{aligned} \text{The Loss function for } G: L^{(G)} = & \min[\log(D(y)) \\ & + \log(1 - D(G(\theta)))]. \end{aligned} \quad (29)$$

Combining equations (28) and (29), the total loss of the entire dataset represented by the min-max value function is given by

$$\begin{aligned} \min_G \max_D V(D, G) = & \min_G \max_D (E_{y \sim P_d(y)} [\log(D(y))] \\ & + E_{\theta \sim P_\theta(\theta)} [\log(1 - D(G(\theta)))]). \end{aligned} \quad (30)$$

GAN is a generative modeling neural network architecture based on the concept of adversarial training that utilizes a model to build new instances that are conceivably derived from an existing sample distribution. Hence, GANs are new favorites for classifying HSIs as they compensate for the lack of data problem and classify the data in a pro manner. There are several types of GANs—conditional GAN, vanilla GAN, deep convolutional GAN (simple type); and Pix2Pix GAN, CycleGAN, StackGAN, and InfoGAN (complex type) [142]. These may be very useful for images like HSIs as they can deal with related issues. The research works based on the GAN are listed in Table 11.

4.7. Transfer Learning (TL). It is the most current hot topic in interactive learning, and there are more to it to be explored. It is an approach where information gained is transferred in one or more source tasks and is used to enhance the learning of a similar target task. TL can be represented diagrammatically by Figure 12 and mathematically shown as follows:

Domain, \mathbf{D} , is represented as $\{\mathbf{X}, P(X)\}$, $X = \{x_1, \dots, x_n\}$, $x_i \in \mathbf{X}$; \mathbf{X} denotes the feature space, and $P(X)$ symbolizes the marginal probability of sample data point X [149].

Task \mathbf{T} is depicted as $\{\mathbf{Y}, P(Y|X)\} = \{\mathbf{Y}, \Phi\}$, $Y = \{y_1, \dots, y_n\}$, $y_i \in \mathbf{Y}$; \mathbf{Y} is the label space, Φ is the prognostic objective function, having learned form (feature vector, label) couples, (x_i, y_i) ; $x_i \in \mathbf{X}$, $y_i \in \mathbf{Y}$, and calculated as the conditional probability.

Also, for every feature vector in \mathbf{D} , Φ predicts its corresponding label as $\Phi(x_i) = y_i$.

If \mathbf{D}_S and \mathbf{D}_T be the source and target domains, \mathbf{T}_S and \mathbf{T}_T be the source and target tasks, respectively, with $\mathbf{D}_S \neq \mathbf{D}_T$ and $\mathbf{T}_S \neq \mathbf{T}_T$. TL objectifies to learn $P(Y_T|X_T)$, that is, the target conditional probability distribution in \mathbf{D}_T with knowledge obtained from \mathbf{D}_S and \mathbf{T}_S .

Traditional learning is segregated and solely based on particular tasks, datasets, and different independent models working on them. No information that can be converted from one model to another is preserved, but on the contrary, TL possesses the human-like capability of transferring

TABLE 11: Summary of review of HSI classification using deep learning—GAN.

Year	Method used	Dataset and COA	Research remarks and future scope
2018	Hyperspectral 1D generative adversarial networks (HSGAN) [140]	IP—83.53%	Outperforms CNN, KNN, etc.
2018	3D augmented GAN [143]	SV—93.67%, IP—91.1%, KSC—98.12%	Data augmentation solved the problem of overfitting and improved class accuracy
2019	Conditional GAN with conditional variational AE (CGAN-CVAE) [144]	UP—83.85%, DC Mall—89.36%	Semi-supervised and ensemble prediction technique ensures the model's training under limited sample conditions
2020	Semi-supervised variational GAN (SSVGAN) [145]	UP—84.35%, Pavia Center—97.15%, DC Mall—92.21%, Jiamusi—64.76%	Outperforms other GAN variants, that is, CVAEGAN and ACGAN, but it suffers from feature matching, overfitting, and convergence problem Correction through metric learning method
2020	Spectral-spatial GAN-conditional random field (SS-GANCRF) [146]	IP—96.3%, UP—99.31%	Enhanced classification capability Creating an end-to-end training system, graph constraint placed on the convolutional layers
2021	Adaptive weighting feature-fusion generative adversarial network (AWF ² -GAN) [147]	IP—97.53%, UP—98.68%	Exploration of the entire joint feature space and fusion of them, joint loss function, and the central loss gained intraclass sensitivity from local neighboring areas and offered an efficient spatial regularization outcome
2021	Variational generative adversarial network with crossed spatial and spectral interactions (CSSVGAN) [148]	IP—93.61%, UP—99.11%, SV—97%	Increased classification potential by utilizing transformer and GAN

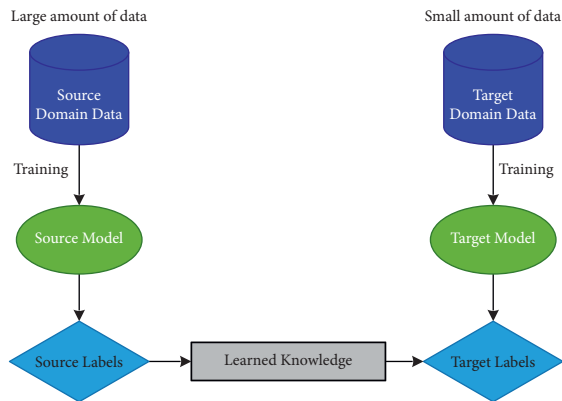


FIGURE 12: The principle of transfer learning.

knowledge; that is, knowledge can be leveraged from priorly trained models to train new models, the process of which is faster, more accurate, and with the limited amount of training data. Table 12 represents a brief detail about the research works on transfer learning.

5. Discussion

Based on the reviewed articles, we can draw the desired inferences that provide answers to the investigative questions mentioned in Section 2 and show the clear motive and benefits of this review.

RI 1: What is the significance of traditional ML and DL for analyzing HSI?

Ans: Hyperspectral data have certain restrictions, as cited in Section 1. Statistical classifiers initially addressed

them, but the operations and analysis became much easier and more accurate after the invention of ML/DL strategies in a machine-dependent way [155, 156]. The general advantages that researchers were provided by the ML/DL algorithms while dealing with HSIs are as follows: (i) easy dealing with high-dimensional data, that is, troubles of Hughes phenomenon removed [115, 125]; (ii) equally manipulative to labeled and unlabeled samples [99, 150]; (iii) precise and the meticulous choice of features [51, 127]; (iv) high-end-precise models to deal with real hypercubes, hence top-notch classification accuracy [119, 154]; (v) removes overfitting, noises, and other hurdles to a much greater extent [120, 147]; (vi) embedded spatial-spectral feature extraction and selection units [119, 133]; (vii) mimics human brain to solve multiclass problems [136, 138].

RI 2: How are ML/DL more impactful on HSI than other non-ML strategies?

Ans: The initial discovery of hyperspectral data has suffered due to its limitations. In the preliminary research stage, the scientists followed the traditional methodology for classifying HSIs, that is, preprocessing (if required), extraction, and selection of discriminative characteristics and then ran a classifier on those features to identify the land cover groups. Hence, they emphasized the feature extractor techniques such as PCA [9], ICA [10], and wavelets [13], assisted by some basic random classifiers such as extended morphological profiles [2, 157], NN [158, 159], logistic regression [160], edge-preserving filters [10, 161], density functions/matrices [162], and Bayes law of classification [163, 164]. These classic mathematics-oriented techniques were not enough to deal with such a huge amount of data like HSI, as they were simple in structure and design and easy to

TABLE 12: Summary of review of HSI classification using transfer learning.

Year	Method used	Dataset and COA	Research remarks and future scope
2018	Deep mapping-based heterogeneous transfer learning model (DLTM) [150]	Washington DC Mall—96.25%	Capable of binary classification Improvisation to multiclass classification
2018	AL with stacked sparse autoencoder (AL-SSAE) [151]	UP—99.48%, center of Pavia—99.8%, SV— 99.45%	Domains, both source, and target possess finely tuned hyperparameters Architectural parameters need to be modified further to enhance the classification accuracy
2020	Heterogeneous TL based on CNN with attention mechanism (HT-CNN-attention) [152]	SV—99%, UP—97.78%, KSC—99.56%, IP—96.99%	Efficient approach regardless of the sample selection strategies chosen
2020	ELM-based ensemble transfer learning (TL-ELM) [26]	UP—98.12%, Pavia center—96.25%	Efficient accuracy and transferability with high training speed Inclusion of SuperPCA and knowledge transfer
2020	Lightweight shuffled group convolutional neural network (abbreviated as SG-CNN) [153]	Botswana—99.67%, HU—99.4%, Washington DC—97.06%	Fine-tuned model as compared to CNN architectures, low computational cost for training Inclusion of more grouped convolutional architectures
2021	Super-pixel pooling convolutional neural network with transfer learning (SP-CNN) [154]	SV—95.99%, UP—93.18%, IP—94.45%	More excellent parameter optimization with more accuracy using a limited number of samples and in a very short period for both training and testing Optimal super-pixel segmentation and merging with different CNN architectures

implement. It also could not predict well enough the multiclass problems, which is very much required for a dataset like HSI, whose land covers belong to multiple classes of regions. Also, these methods were not accurate in feature selection and extraction or dealing with the storage of such bulk data. These reasons made researchers struggle to analyze properly, process, and classify HSIs. On the contrary, the advancements of ML/DL technologies have opened a broad gateway of research that researchers are still exploring and combining with different groupings to address the HSI classification problem in real life, dealing with the limitations mentioned above [26, 131]. The tabular depiction of the advantages and disadvantages of the ML and non-ML strategies applied for HSI classification is shown in Table 13.

RI 3: What are the advantages and challenges faced by the researchers for the chosen ML/DL-based algorithm for HSI classification?

Ans: We added the advantages and challenges of the ML- and DL-based techniques in Table 13.

RI 4: What are the emerging literary works of ML/DL on HSI classification in the year 2021?

Ans: In the ongoing years, 2021 seems to be more promising in terms of technical advancements for the problem concerned. New techniques are emerging, along with hybrid ones, to solve the issue to a whole new level, the methodologies' accuracy to be described. Recent work on MRF with a band-weighted discrete spectral mixture model (MRF-BDSMM) in a Bayesian framework has been proposed in [165], an unsupervised adaptive approach to accommodate heterogeneous noise and find the abundant labeled subpixels to extricate joint features. A collaboration of Kernel-based ELM with PCA, local binary pattern (LBP), and gray-wolf optimization algorithm (PLG) is proposed as novel methodologies. They help reduce huge dimensions, seek global and local-spatial features, and

optimize the KELM parameters to obtain the class labels [166]. A variant of SRC is proposed in [167], dual sparse representation graph-based collaborative propagation (DSRG-CP) that separates spatial and spectral dimensions with the respective graph to improve the labeling scheme limited samples by collaborating the outcomes. AL has been one of the hot topics so far, as it integrates with a Fredholm kernel regularized model (AMKFL) that enables better labeling than manual ones, even for noisy images [168]. It ties with DL with the augmentation of training samples to label the uncertain hypercubes (ADL-UL) accurately [169], facilitates iterative training sample augmentation by expanding the hypercubes and adds discriminative joint features (ITSA-AL-SS) [170], extracts local unique spatial multi-scale characteristics from the super-pixels (MSAL) [171]. A novel idea of attention-based CNNs is proposed in [172, 173], the former (SSAtt-CNN) collides two attention subnetworks—spatial and spectral with CNN as the base, and the latter (FADCNN) is a dense spectral-spatial CNN with feedback attention technique that perfectly poses the band weights for better mining and utilization of dominant features. GAN is one the most exploited methods to date, and [174] proposes the full utilization of shallow features from the unlabeled bands through a multitasking network (MTGAN); in [175], the discriminator is based upon capsule network and convolutional long short-term memory to extricate less visible features and integrates them to build high-profile contextual characteristics (CCAPS-GAN); 1D and 2D CapsGAN together form a dual-channel spectral-spatial fusion capsule GAN (DcCaps-GAN) shown in [176]; and generative adversarial minority oversampling for 3D-hypercubes (3D-HyperGAMO) is depicted in [177] that focuses on the minor class features using existing ones to label and classify them properly.

TABLE 13: Comparison between ML and non-ML techniques for HSI classification.

Methods	Advantages	Disadvantages
Classical state-of-art techniques	<ul style="list-style-type: none"> (i) Simple structure and design (ii) Less time consumption (iii) Easy to implement (iv) Dimension handling skillfully by PCA and ICA (v) Better binary and moderate multiclass classification by kernel and SVM 	<ul style="list-style-type: none"> (i) High space complexity due to the storage of bulk data (ii) Based on empirical identities, hence a tedious workpiece (iii) Feature selection and extraction are not accurate (iv) Suffers from limited labeled sample problem, Hughes phenomenon, and noise
Advanced machine learning techniques	<ul style="list-style-type: none"> (i) Easy dealing with high-dimensional data, that is, troubles of Hughes phenomenon removed (ii) Equally manipulative to labeled and unlabeled samples (iii) Precise and meticulous choice of features (iv) High-end-precise models to deal with real hypercubes, hence, top-notch classification accuracy (v) Removes overfitting, noises, and other hurdles to a much greater extent (vi) Mimics the human brain to solve multiclass problems 	<ul style="list-style-type: none"> (i) The construction of the model is difficult due to its complex network-alike structure (ii) High time complexity due to training and testing of the huge amount of raw HSI data (iii) Extremely expensive design (iv) Strenuous to implement

RI 5: How are ML- and DL-based hybrid techniques helping scientists in HSI classification?

Ans: Since the dawn of the emergence of HSIs, it has suffered many hurdles in its path of analysis and information extraction. The maximum number of highly correlated bands and the high spatial-spectral features signature by the electromagnetic spectrum embedded in it are always considered a traction matter. Thus, finding an appropriate technology for the classification of such interconnected and hugely confined featured high-dimensional images is a very tedious and strenuous matter. The classification methods chosen so far have been mostly limited to supervised. The requirement of a sufficient number of quality-labeled data and unsupervised, in which the lack of coherence between the spectral clusters and the target regions, causes the failure in obtaining the desired accuracy. A semi-supervised method is needed to overcome such problems as a combination of supervised and unsupervised methods, named the hybrid method. A hybrid method is always advantageous in robustness and flexibility towards the high-dimensional data.

The hybrid methods have the following benefits:

- (i) Specifically designed to overcome the limitations and take advantage of the methodologies involved in the concerned hybrid to achieve a deep, rich, and insightful conclusion (general).
- (ii) Addressing and resolving multiple issues regarding the handling and analyzing the HSI data, at a time, depending upon the methods that are chosen for mixing/hybridizing [179–183].
- (iii) Coherence in time, space, and cost complexities [184–186].
- (iv) Better interpretability, quality, effectivity leading to the construction of a more refined framework [180, 182, 183, 187–194].

- (v) Deterministic spectral, spatial, and contextual feature extraction, reduction, and selection, and combining them to achieve desired accuracy and performance [182, 183, 187, 188, 195–197].

ML, being a standard versatile technology, can merge with traditional techniques like PCA for its benefit. As stated in [195, 198], PCA is exploited at its best for feature extraction, selection, and reduction to achieve higher accuracy and performance quality. PCA is one of the best pre-processing methods considered to date for improvised spectral dimension reduction [180], proper selection of spectral bands and their multiscale features in a segmented format [181, 199], noise-reduced spectral analysis [27], and feature extraction [130, 196]. PCA, in collaboration with SVM [195, 200], DL for feature reduction and better classification [182, 183], CNN with multiscale feature extraction [188, 189], and sparse tensor technology [190], has highly been appreciated as soulful research. All these recent time collaborations and a special honor to the merging of ICA-DCT with CNN cited in [191] are the evidence that although PCA is categorized under traditional methods, it is supremely relevant for its significant usefulness in handling HSIs.

Some other hybridizations are also explored by researchers, such as SRC with mathematical index of divergence-correlation [192], Gabor-cube filter [193], and ELM [83, 85]; ELM with CNN [86] and TL [26]; AL based on super-pixel profile [201, 202], AL with CNN [203], CapsNet [204], CNN [204, 205], and TL [151, 184]; CNN with attention-aided methodology [172, 173, 185] and GAN [186]; GAN with dynamic neighborhood majority voting mechanism [194, 197], CapsNet [175, 176, 206, 207]; and TL with MRF [70]. These articles depict the highly tenacious performance with literal mitigation of the computational complexities enforced on the raw HSI data to build a strong and enhanced model for achieving higher accuracy than ever.

TABLE 14: The advantages and challenges of the ML- and DL-based techniques for HSI classification.

ML/DL techniques	Advantages	Challenges
Support vector machine	<ul style="list-style-type: none"> (i) Robust in terms of outliers, Hughes effect, and dimensions as its reduction is not primarily necessary [32, 41, 43] (ii) Supports both supervised, semi-supervised, and unsupervised problems with less overfitting risks [24, 33, 37, 44] (iii) Form of a sigmoid kernel that deals better than the rest of the previous for unlabeled and unstructured HSI datasets [35, 40–42] (iv) The capability of solving the classification problem for both binary and multiclass problems by outperforming several methods [39] (v) Can improve the performance if assisted with other supporting methods [36, 40–42] 	<ul style="list-style-type: none"> (i) It works very well for binary classification but fails for generating accurate classes for multiclass problems [31] (ii) Training time is high for high-class datasets like HSI [31, 32] (iii) Difficulty in fine-tuning the parameters [41, 42] (iv) Complex interpretability [33, 35] (v) Lack of easy generalization to the datasets having multiple classes [33, 35] (vi) Complexity in building the model due to a lack of sufficient labeled samples [31, 32]
Sparse representation and classification	<ul style="list-style-type: none"> (i) A dictionary with relevant data is used for learning with a minimal number of optimal parameters [45, 46] (ii) Builds precise and powerful classification models with higher interpretability through sparse coding [49, 50, 54] (iii) Proper memory usage in an optimized manner [53, 55, 178] (iv) Reduces the estimated variance between the classes to produce better outcomes [49, 56, 178] 	<ul style="list-style-type: none"> (i) Making the dictionary considers high expense overheads [50] (ii) The dictionary or the coding might cause loss of information [48, 178] (iii) Difficulties in representing such high-profile with higher resolution image data like HSI through the sparse matrix [47, 48]
Markov random field	<ul style="list-style-type: none"> (i) Works well for a wide range of unstructured problems and no direct dependency between classes and the parameters [67, 69] (ii) Better denoising effect [59] (iii) Robust for both spatial and spectral distributions [62, 64] (iv) Time complexity is low due to the graphical representation of data [63] 	<ul style="list-style-type: none"> (i) Normalization of data might be hectic for high dimension data [63, 70] (ii) Suffers from the lack of training undirected data that might not be possible to represent graphically [61, 62] (iii) Poor interpretability [63, 68]
Extreme learning machines	<ul style="list-style-type: none"> (i) Less training time and faster learning rate as compared to previous methods [86] (ii) Avoidance for local minima and finishes job in single iteration [83, 87] (iii) Advantageous for overfitting caused due to several bands in HSIs [83] (iv) Builds an enhanced model with better prediction performance at the optimized expense [86] (v) Improved generalization ability, robustness, and controllability [78, 84, 85] 	<ul style="list-style-type: none"> (i) Higher computational hazard [76–80] (ii) The wrong choice of an optimal amount of the hidden layer neurons may cause redundancy in the model and hence affect the classification accuracy [85, 86] (iii) There is plenty of room for advancements in the algorithm to accommodate itself to be compatible for dealing with HSI data [78, 82, 86]

TABLE 14: Continued.

ML/DL techniques	Advantages	Challenges
Active learning	(i) A very efficient way of learning for both supervised and semi-supervised problems [91, 97, 101, 103]	(i) Higher computational hazard [76–80]
	(ii) Ease in segregating the interclass and intraclass features through active query sets [91, 95, 102, 103]	(ii) The wrong choice of an optimal amount of the hidden layer neurons may cause redundancy in the model and hence affect the classification accuracy [85, 86]
	(iii) Training speed is comparatively high for not so large-scale data [103]	(iii) There is plenty of room for advancements in the algorithm to accommodate itself to be compatible for dealing with HSI data [78, 82, 86]
	(iv) Knowledge-based solid models can be generated [103]	
	(v) Achieves greater classification accuracies for unlabeled HSIs [95, 102]	
Deep learning	(i) Diverse, unstructured, and unlabeled raw HSI datasets are finely processed where preprocessing of the data is not needed [110, 122, 125, 144]	(i) Suffers from a lack of a large amount of HSI data, which is practically unavailable [123, 136]
	(ii) Possesses the capability to address supervised, semi-supervised, and specifically unsupervised learning problems [127, 128, 137]	(ii) The extreme expense to generate an appropriate model by training a complex data structure like HSIs [114, 139, 148]
	(iii) Expertise in dimension reduction, denoising, feature extraction as embedded properties [27, 114, 124]	(iii) Low interpretability [131, 147]
	(iv) Address in an illustrious manner to the issues such as Hughes phenomenon, overfitting, and convergence. [120, 124, 145]	(iv) Theoretically not sound, hence incomprehensible where an error occurs and its rectification [122, 124, 145]
	(v) Robust and adaptive to new features introduced in the dataset [26, 123, 145]	(v) High time and space complexity and computational hazard [131, 136, 148]
	(vi) The hidden layer neurons are proven to be eminent in training the desired model with a highly qualified prior knowledge (DBN, RNN, CNN) [127, 129, 135, 138]	
	(vii) Computational efficiency with high-performance speed (CNN, SAE) [114, 115, 127, 128]	
	(viii) Data augmentation facility (GAN) [143, 145]	
Transfer learning	(i) Works as a combination of different models, be it traditional or latest machine-learned techniques, that together brings out a highly improved hybrid model [151, 152]	(i) Data overfitting [150]
	(ii) Capable of transferring knowledge from the source domain, that is, a pretrained model to the target domain, that is, a new model to make it more enriched [151, 152]	(ii) Complex structure of the model [150, 151]
	(iii) Greater feature extraction and selection capability [152]	(iii) Less interpretability
	(iv) Stable model with highly optimized parameters and hyperparameters [154]	(iv) Difficulty in implementation
	(v) High training speed and accuracy with low computational cost [26, 153]	
	(vi) Reduced computational cost and training time complexity [153, 154]	

RI 6: What are the latest emerging techniques associated with addressing classifying HSIs?

Ans: The following are the most recent research studies that have enlightened a new path of dealing with the purpose:

- (i) *DSVM*: The latest and novel concept incorporates DL facilities with traditional kernel SVM. This combines four deep layers of kernels with SVM being the hidden layer units, namely, exponential and gaussian radial basis function (ERBF and

GRBF), neural and polynomial [208]. This approach has outperformed several efficient DL methods with nearly 100% accuracy for IP and UP datasets.

- (ii) *Conditional Random Fields (CRFs)*: These are the structured generalization of multinomial logistic regression in the form of graphical models based on a priori continuity considering the neighboring pixels of analogous spectral signatures that possess the same labels. They extensively explore the hidden

spectral-contextual information. In [146], CRF incorporates with semi-supervised GAN whose trained discriminators produce softmax predictions that are guided by dense CRFs graph constraints to improve HSI classification maps. A collaboration between 3D-CNN and CRF has been proposed in [209] to make a deep CRF capable of extracting the semantic correlations between patches of hypercubes by CNN's unary and pairwise potential functions. A semi-supervised approach is depicted in [210], embedding subspace learning and 3D convolutional autoencoder to remove redundancy in joint features and obtain class sets using an iterative algorithm. In [211], CRF with Gaussian edge potentials associated with deep metric learning (DML) classifies HSI data pixelwise using the geographical distances between pixels and the Euclidean distances between the features. A novel framework using HSI feature learning network (HSINet) with CRF is proposed [212] that is a trainable end-to-end DL model with back-propagation that extracts joint features, edges, and colors based on subpixel, pixel, and super-pixels. In [213], a decision fusion model including CRF and MRF is built based on sparse unmixing and soft classifiers output.

- (iii) *Random Forest (RF)*: It is an efficient algorithm that ensembles regression and classification tree. It enables the HSI classification model to be noise-tolerant, inherent in the multiclass division, robustness in parallelism, and speed. In [214], RF is compared to the DL algorithm, which outshined the classification accuracy. A new framework of cascaded RF is shown in [215] that uses the boosting strategy to generate and train base classifiers and Hierarchical Random Subspace Method to select features and suitable base classifiers based on the diversity of the features. A novel collaboration of semi-supervised learning and AL and RF is featured in [216], where the queries based on spatial information are fed to AL, and then, the labeled samples are classified by RF through semi-supervision. [217, 218] depicts a deep cube CNN model that extracts pixelwise joint features and is classified by RF.
- (iv) *Graph Convolutional Network (GCN)*: A descendent of CNN, a structure designed to generalize and convert the convolution data to graph data. It consists of three steps feature aggregation, feature transformation, and classification. Being an expert in graphical modeling considers the spatial interrelations between the classes at its best. In [219], the different unique features collected from CNN and GCN are fused additive, elementwise, and concatenated way. A new framework of globally consistent GCN is introduced in [220], which first generates a spatial-spectral local optimized graph whose global high-order neighbors obtain the enriched contextual information employing the graph topological

consistent connectivity; at last, those global features determine the classes. [221] shows the concept of a dual GCN network, which works with a limited number of training samples, where first extricates all the significant features and second learns label distribution. A novel idea of deep attention GCN is introduced in [222] based on similarity measurement criteria between the mixed measurement of a kernel-spectral angle mapper and spectral information divergence to accumulate analogous spectra. [223] emerges as a collaboration between CNN and GCN to extract pixel and super-pixelwise joint features by learning small-scale regular regions and large-scale irregular regions.

6. Conclusion

This article depicts the various technologies and procedures used for HSI classification since the dawn of its invention to date. There are many barriers to dealing with such high-band data as HSI mentioned above. Despite that, many researchers have taken their interest in this field to improve the existing techniques or even invent new ones throughout the last decade. As per the considerable improvement in technologies and the introduction of ML into the classification issues of HSI, it has become more accurate than traditional and contemporary state-of-art methodologies. As a result, DL has emerged as the most eminent work tool for HSI classification for the last half of this decade. The more the researchers focused on this, the more they explored the remote sensing and space imagery features.

This review article bears the individual information for every method and their submethods about their performance, research gaps, and achievements. In addition, it appends a novel research methodology that makes this work more distinctive than others. After going through each methodology's minute details, the most significant inferences have been drawn, which add further novelty to our work. Also, it shows a path of choosing an appropriate technique and its alternatives for future researchers, hence alleviating its creativity and uniqueness, above all other contemporary review works on this subject. Also, it provides the details of the most recent research scenario on HSI classification and some of the currently developed techniques that might be acutely useful in several future research. Our study holds the uniqueness and the novelty regarding several aspects, such as the following: (1) it includes the research works carried out in the last decade, that is, 2010–2020, and the most recent papers of the previous year, i.e., 2021, and we have mentioned it in Section 3; (2) the number of papers referred here is above 200, outnumbering other review papers; (3) the review is carried out by selecting the most appropriate papers solely dedicated to our subject of interest, that is, machine learning techniques serving the purpose of hyperspectral image classification. Then, the findings from those works of literature are systematically arranged in the tabular format (Tables 1–12); (4) the

objective behind this review work is expressed by RQ 1–6. Also, they provide a clear view of the recent technological advances and applications that the researchers are developing in recent times; (5) Table 14 provides an explicit idea of the pros and cons of each ML technique described in this manuscript when applied for classifying hyperspectral images, which will help the researchers in their future research; and (6) the researcher who wishes to write a literature review can follow our proposed methodology that depicts the flow of work in a methodical way. [224].

7. Limitations of Present Work and Its Future Scope

The study has some limitations: (i) we have used fewer keywords in the current research (ii) we only focused on seven popular ML techniques; (iii) we briefly explain the emerging methodologies; and (iv) the experimental details are not fully discussed.

As a future proposition, we would like to explore more keywords, more techniques, and more studies that offer a better understanding of other learning methods, both traditional and contemporary. In addition, there are several instances of hybrid strategies along with some more eminent and latest ML/DL techniques that we shall look forward to exploring in both qualitative and quantitative manner.

Acronym

HS:	Hyperspectral
HSI:	Hyperspectral image
GIS:	Geographic Information System
PCA:	Principal component analysis
ICA:	Independent component analysis
SVM:	Support vector machine
SR:	Sparse representation
SRC:	Sparse representation and classification
MRF:	Markov random field
HMRF:	Hidden Markov random field
ELM:	Extreme learning machine
AL:	Active learning
HU:	University of Houston
TL:	Transfer learning
DL:	Deep learning
AE:	Autoencoders
SAE:	Stacked autoencoders
CNN:	Convolutional neural network
RNN:	Recurrent neural network
DBN:	Deep belief network
GAN:	Generative adversarial network
IP:	Indian pines
KSC:	Kennedy space center
SV:	Salinas valley
UP:	University of Pavia.

Data Availability

Publicly available data are used in this study.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

Jana Shafi would like to thank the Deanship of Scientific Research, Prince Sattam bin Abdul Aziz University, for supporting this work. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (Grant no. 2022R1C1C1004590).

References

- [1] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: a review," *IEEE Access*, vol. 6, pp. 14118–14129, 2018.
- [2] N. Falco, J. A. Benediktsson, and L. Bruzzone, "Spectral and spatial classification of hyperspectral images based on ICA and reduced morphological attribute profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 6223–6240, 2015.
- [3] T. Adão, J. Hruška, L. Pádua et al., "Hyperspectral imaging: a review on UAV-based sensors, data processing and applications for agriculture and forestry," *Remote Sensing*, vol. 9, p. 1110, 2017.
- [4] C. Northcutt, L. Jiang, and I. Chuang, "Confident learning: estimating uncertainty in dataset labels," <https://www.jair.org/index.php/jair/article/view/12125>.
- [5] Z. Xu, D. Lu, Y. Wang et al., "Noisy labels are treasure: mean-teacher-assisted confident learning for hepatic vessel segmentation," in *Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Springer, Strasbourg, France, 27 September 2021, https://link.springer.com/chapter/10.1007%2F978-3-030-87193-2_1.
- [6] C. G. Northcutt, T. Wu, and I. L. Chuang, "Learning with confident examples: rank pruning for robust classification with noisy labels," <https://arxiv.org/abs/1705.01936>.
- [7] P. Ghamisi, N. Yokoya, J. Li et al., "Advances in hyperspectral image and signal processing: a comprehensive overview of the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 37–78, Dec. 2017.
- [8] T. Han and D. G. Goodenough, "Investigation of nonlinearity in hyperspectral imagery using surrogate data methods," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 10, pp. 2840–2847, Oct. 2008.
- [9] B. A. Beirami and M. Mokhtarzade, "Band grouping SuperPCA for feature extraction and extended morphological profile production from hyperspectral images," *IEEE Geoscience and Remote Sensing Letters (Early Access)*, vol. 17, pp. 1–5, 2020.
- [10] J. Xia, L. Bombrun, T. Adali, Y. Berthoumieu, and C. Germain, "Spectral–spatial classification of hyperspectral images using ICA and edge-preserving filter via an ensemble strategy," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4971–4982, 2016.
- [11] M. Imani and H. Ghassemian, "Principal component discriminant analysis for feature extraction and classification of hyperspectral images," in *Proceedings of the 2014 Iranian Conference on Intelligent Systems (ICIS)*, IEEE, Bam, Iran, 4 February 2014.

- [12] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving discriminant analysis in kernel-induced feature spaces for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 5, pp. 894–898, 2011.
- [13] X. Cao, J. Yao, X. Fu, H. Bi, and D. Hong, "An enhanced 3-D discrete wavelet transform for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters (Early Access)*, vol. 18, pp. 1–5, 2020.
- [14] J. Peng, H. Chen, Y. Zhou, and L. Li, "Ideal regularized composite kernel for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 4, pp. 1563–1574, 2017.
- [15] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 9, pp. 4816–4829, 2013.
- [16] J. Liu, Z. Wu, J. Li, A. Plaza, and Y. Yuan, "Probabilistic-kernel collaborative representation for spatial-spectral hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 2371–2384, 2016.
- [17] M. S. Kumar, V. Keerthi, R. N. Anjani, M. M. Sarma, and V. Bothale, "Evaluation of machine learning methods for hyperspectral image classification," in *Proceedings of the 2020 IEEE India Geoscience and Remote Sensing Symposium (InGARSS)*, pp. 225–228, IEEE, Ahmedabad, India, 1 December 2020.
- [18] "Hyperspectral remote sensing scenes," http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes.
- [19] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: a review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 158, pp. 279–317, December, 2019.
- [20] C. L. Chowdhary, P. V. Patel, K. J. Kathrotia, M. Attique, K. Perumal, and M. F. Ijaz, "Analytical study of hybrid techniques for image encryption and decryption," *Sensors*, vol. 20, no. 18, p. 5162, 2020.
- [21] S. Jia, S. Jiang, Z. Lin, N. Li, M. Xu, and S. Yu, "A survey: deep learning for hyperspectral image classification with few labeled samples," *Neurocomputing*, vol. 448, pp. 179–204, 2021.
- [22] Y. Quan, X. Zhong, W. Feng, J. C.-W. Chan, Q. Li, and M. Xing, "SMOTE-based weighted deep rotation forest for the imbalanced hyperspectral data classification," *Remote Sensing*, vol. 13, no. 3, p. 464, 2021.
- [23] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, no. 1, pp. 55–63, January 1968.
- [24] D. K. Pathak and S. K. Kalita, "Spectral spatial feature based classification of hyperspectral image using support vector machine," in *Proceedings of the 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), Date of Conference*, IEEE, Noida, India, 7 March 2019.
- [25] L. Zhou and L. Ma, "Extreme learning machine-based heterogeneous domain adaptation for classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 11, pp. 1781–1785, 2019.
- [26] X. Liu, Q. Hu, Y. Cai, and Z. Cai, "Extreme learning machine-based ensemble transfer learning for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3892–3902.
- [27] H. Chen, F. Miao, and X. Shen, "Hyperspectral remote sensing image classification with CNN based on quantum genetic-optimized sparse representation," *IEEE Access*, vol. 8, pp. 99900–99909.
- [28] H. C. Li, S. S. Li, W. S. Hu, J. H. Feng, W. W. Sun, and Q. Du, "Recurrent feedback convolutional neural network for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, 2021.
- [29] S. Mian Qaisar, "Signal-piloted processing and machine learning based efficient power quality disturbances recognition," *PLoS One*, vol. 16, no. 5, Article ID e0252104, 2021.
- [30] F. Chamasemani and Y. P. Singh, "Multi-class support vector machine (SVM) classifiers -- an application in hypothyroid detection and classification," in *Proceedings of the The 2011 Sixth International Conference on Bio-Inspired Computing*, pp. 351–356, IEEE, Penang, Malaysia, 27 September 2011.
- [31] J. Zhang, Y. Zhang, and T. Zhou, "Classification of hyperspectral data using support vector machine," in *Proceedings of the 2001 International Conference on Image Processing (Cat. No.01CH37205)*, IEEE, Thessaloniki, Greece, 7 October 2001.
- [32] G. Camps-Valls, L. Gomez-Chova, J. Calpe-Maravilla et al., "Robust support vector method for hyperspectral data classification and knowledge discovery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 7, pp. 1530–1542, July 2004.
- [33] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1778–1790, 2004.
- [34] B. Borasca, L. Bruzzone, L. Carlin, and M. Zusi, "A fuzzy-input fuzzy-output SVM technique for classification of hyperspectral remote sensing images," in *Proceedings of the 7th Nordic Signal Processing Symposium - NORSIG 2006*, IEEE, Reykjavik, Iceland, 7 June 2006.
- [35] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3804–3814, 2008.
- [36] S. Ding and L. Chen, "Classification of hyperspectral remote sensing images with support vector machines and particle swarm optimization," in *Proceedings of the International Conference on Information Engineering and Computer Science*, 19 December 2009.
- [37] P. Du, K. Tan, and X. Xing, "Wavelet SVM in Reproducing Kernel Hilbert Space for hyperspectral remote sensing image classification," *Optics Communications*, vol. 283, no. 24, pp. 4978–4984, 2010.
- [38] F. A. Mianji and Y. Zhang, "Semisupervised support vector machine classification for hyperspectral imagery," in *Proceedings of the 2011 International Conference on Communications and Signal Processing*, 10 February 2011.
- [39] S. Moustakidis, G. Mallinis, N. Koutsias, J. B. Theodoris, and V. Petridis, "SVM-based fuzzy decision trees for classification of high spatial resolution remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 149–169, Jan. 2012.
- [40] Z. Shao, L. Zhang, X. Zhou, and L. Ding, "A novel hierarchical semisupervised SVM for classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 9, pp. 1609–1613, 2014.

- [41] B. Kuo, H. Ho, C. Li, C. Hung, and J. Taur, "A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 1, pp. 317–326, Jan. 2014.
- [42] J. Peng, Y. Zhou, and C. L. P. Chen, "Region-kernel-based support vector machines for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 9, pp. 4810–4824, 2015.
- [43] H. Yu, L. Gao, W. Liao, B. Zhang, A. Pižurica, and W. Philips, "Multiscale super-pixel-level subspace-based support vector machines for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 2142–2146, 2017.
- [44] C. Zhang, M. Han, and M. Xu, "Multi-feature classification of hyperspectral image via probabilistic SVM and guided filter," in *Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN)*, IEEE, Rio de Janeiro, Brazil, 8 July 2018.
- [45] J. Liu, Z. Wu, Z. Wei, L. Xiao, and L. Sun, "Spatial-spectral kernel sparse representation for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 6, pp. 2462–2471, Dec. 2013.
- [46] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral-spatial hyperspectral image classification via multiscale Adaptive sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 12, pp. 7738–7749, Dec. 2014.
- [47] P. Du, Z. Xue, J. Li, and A. Plaza, "Learning discriminative sparse representations for hyperspectral image classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 1089–1104, Sept. 2015.
- [48] W. Fu, S. Li, L. Fang, X. Kang, and J. A. Benediktsson, "Hyperspectral image classification via shape-adaptive joint sparse representation," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 2, pp. 556–567, Feb. 2016.
- [49] L. Fang, C. Wang, S. Li, and J. A. Benediktsson, "Hyperspectral image classification via multiple-feature-based adaptive sparse representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1646–1657, July 2017.
- [50] B. Tu, S. Huang, L. Fang, G. Zhang, J. Wang, and B. Zheng, "Hyperspectral image classification via weighted joint nearest neighbor and sparse representation," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 4063–4075, 2018.
- [51] W. Yang, J. Peng, W. Sun, and Q. Du, "Log-euclidean kernel-based joint sparse representation for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 12, pp. 5023–5034, Dec. 2019.
- [52] T. Dundar and T. Ince, "Sparse representation-based hyperspectral image classification using multiscale super-pixels and guided filter," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 2, pp. 246–250, Feb. 2019.
- [53] J. Peng, W. Sun, and Q. Du, "Self-paced joint sparse representation for the classification of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 1183–1194, Feb. 2019.
- [54] J. Peng, L. Li, and Y. Y. Tang, "Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 6, pp. 1790–1802, June 2019.
- [55] H. Yu, L. Gao, W. Liao et al., "Global spatial and local spectral similarity-based manifold learning group sparse representation for hyperspectral imagery classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3043–3056, May 2020.
- [56] F. Luo, L. Zhang, X. Zhou, T. Guo, Y. Cheng, and T. Yin, "Sparse-adaptive hypergraph discriminant analysis for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 6, pp. 1082–1086, June 2020.
- [57] "Markov random field," https://en.wikipedia.org/wiki/Markov_random_field.
- [58] G. Altalib and E. Ahmed, "Land cover classification using hidden Markov models," *International Journal of Computer Networks and Communications Security*, vol. 1, pp. 165–172, 2013.
- [59] B. Zhang, S. Li, X. Jia, L. Gao, and M. Peng, "Adaptive Markov random field approach for classification of hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 5, pp. 973–977, 2011.
- [60] P. Ghamisi, J. A. Benediktsson, and M. O. Ulfarsson, "Spectral-spatial classification of hyperspectral images based on hidden Markov random fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2565–2574, May 2014.
- [61] L. Xu and J. Li, "Bayesian classification of hyperspectral imagery based on probabilistic sparse representation and Markov random field," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 4, pp. 823–827, April 2014.
- [62] W. Li, S. Prasad, and J. E. Fowler, "Hyperspectral image classification using Gaussian mixture models and Markov random fields," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 1, pp. 153–157, Jan. 2014.
- [63] L. Sun, Z. Wu, J. Liu, L. Xiao, and Z. Wei, "Supervised spectral-spatial hyperspectral image classification with weighted Markov random fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1490–1503, March 2015.
- [64] Y. Yuan, J. Lin, and Q. Wang, "Hyperspectral image classification via multitask joint sparse representation and stepwise MRF optimization," *IEEE Transactions on Cybernetics*, vol. 46, no. 12, pp. 2966–2977, Dec. 2016.
- [65] M. Golipour, H. Ghassemian, and F. Mirzapour, "Integrating hierarchical segmentation maps with MRF prior for classification of hyperspectral images in a bayesian framework," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 805–816, Feb. 2016.
- [66] E. K. Ghasrodashti, M. S. Helfroush, and H. Danyali, "Sparse-based classification of hyper-spectral images using extended hidden Markov random fields," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 4101–4112, 2018.
- [67] Y. Fang, L. Xu, J. Peng, H. Yang, A. Wong, and D. A. Clausi, "Unsupervised bayesian classification of a hyperspectral image based on the spectral mixture model and Markov random field," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 9, pp. 3325–3337, 2018.
- [68] C. Pan, X. Gao, Y. Wang, and J. Li, "Markov random fields integrating adaptive interclass-pair penalty and spectral similarity for hyperspectral image classification," *IEEE*

- Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2520–2534, May 2019.
- [69] X. Cao, X. Wang, D. Wang, J. Zhao, and L. Jiao, “Spectral-spatial hyperspectral image classification using cascaded Markov random fields,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 12, pp. 4861–4872, Dec. 2019.
- [70] X. Jiang, Y. Zhang, Y. Li, S. Li, and Y. Zhang, “Hyperspectral image classification with transfer learning and Markov random fields,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 3, pp. 544–548, March 2020.
- [71] X. Jiang, Y. Zhang, W. Liu et al., “Hyperspectral image classification with CapsNet and Markov random fields,” *IEEE Access*, vol. 8, pp. 191956–191968.
- [72] A. Samat, P. Du, S. Liu, J. Li, and L. Cheng, “E2LMs: ensemble extreme learning machines for hyperspectral image classification,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1060–1069, April 2014.
- [73] S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, “Extreme learning machine: algorithm, theory and applications,” *Artificial Intelligence Review*, vol. 44, no. 1, pp. 103–115, 2015.
- [74] “A multiple hidden layers extreme learning machine method and its application,” <https://www.hindawi.com/journals/mpe/2017/4670187/>.
- [75] Y. Zhou, J. Peng, and C. L. P. Chen, “Extreme learning machine with composite kernels for hyperspectral image classification,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2351–2360, June 2015.
- [76] W. Li, C. Chen, H. Su, and Q. Du, “Local binary patterns and extreme learning machine for hyperspectral imagery classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3681–3693, July 2015.
- [77] Q. Lv, X. Niu, Y. Dou, J. Xu, and Y. Lei, “Classification of hyperspectral remote sensing image using hierarchical local-receptive-field-based extreme learning machine,” *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 3, pp. 434–438, March 2016.
- [78] H. Su, Y. Cai, and Q. Du, “Firefly-algorithm-inspired framework with band selection and extreme learning machine for hyperspectral image classification,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 1, pp. 309–320, Jan. 2017.
- [79] Y. Shen, J. Chen, and L. Xiao, “Supervised classification of hyperspectral images using local-receptive-fields-based kernel extreme learning machine,” in *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, Beijing, China, 17–20 September 2017.
- [80] J. Ku and B. Zheng, “Distributed extreme learning machine with kernels based on MapReduce for spectral-spatial classification of hyperspectral image,” in *Proceedings of the 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, IEEE, Guangzhou, China, 21 July 2017.
- [81] F. Cao, Z. Yang, M. Jiang, W. Chen, Q. Ye, and W. Ling, “Spectral-spatial classification of hyperspectral image using extreme learning machine and loopy Belief propagation,” in *Proceedings of the 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, IEEE, Exeter, UK, 21 June 2017.
- [82] W. Shang, Z. Wu, Y. Xu, Y. Zhang, and Z. Wei, “Hyperspectral supervised classification using mean filtering based kernel extreme learning machine,” in *Proceedings of the 2018 Fifth International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*, IEEE, Xi’an, China, 18 June 2018.
- [83] F. Cao, Z. Yang, J. Ren et al., “Sparse representation-based augmented multinomial logistic extreme learning machine with weighted composite features for spectral-spatial classification of hyperspectral images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6263–6279, 2018.
- [84] M. Jiang, F. Cao, and Y. Lu, “Extreme learning machine with enhanced composite feature for spectral-spatial hyperspectral image classification,” *IEEE Access*, vol. 6, pp. 22645–22654.
- [85] F. Cao, Z. Yang, J. Ren, W. Chen, G. Han, and Y. Shen, “Local block multilayer sparse extreme learning machine for effective feature extraction and classification of hyperspectral images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5580–5594, 2019.
- [86] Y. Shen, L. Xiao, J. Chen, and D. Pan, “A spectral-spatial domain-specific convolutional deep extreme learning machine for supervised hyperspectral image classification,” *IEEE Access*, vol. 7, pp. 132240–132252.
- [87] Y. Yin and L. Wei, “Hyperspectral image classification using comprehensive evaluation model of extreme learning machine based on cumulative variation weights,” *IEEE Access*, vol. 8, p. 187991, 188003.
- [88] “Introduction to active learning,” <https://towardsdatascience.com/introduction-to-active-learning-117e0740d7cc>.
- [89] “Active learning machine learning: what it is and how it works,” <https://algorithmia.com/blog/active-learning-machine-learning>.
- [90] S. Rajan, J. Ghosh, and M. M. Crawford, “An active learning approach to hyperspectral data classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 4, pp. 1231–1242, April 2008.
- [91] J. Li, J. M. Bioucas-Dias, and A. Plaza, “Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 11, pp. 4085–4098, 2010.
- [92] J. Li, J. M. Bioucas-Dias, and A. Plaza, “Spectral-spatial classification of hyperspectral data using loopy Belief propagation and active learning,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 2, pp. 844–856, Feb. 2013.
- [93] S. Sun, P. Zhong, H. Xiao, and R. Wang, “An MRF model-based active learning framework for the spectral-spatial classification of hyperspectral imagery,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 1074–1088, Sept. 2015.
- [94] S. Sun, P. Zhong, H. Xiao, and R. Wang, “Active learning with Gaussian process classifier for hyper-spectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 1746–1760, April 2015.
- [95] Z. Zhang, E. Pasolli, M. M. Crawford, and J. C. Tilton, “An active learning framework for hyperspectral image classification using hierarchical segmentation,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 2, pp. 640–654, Feb. 2016.
- [96] X. Zhou, S. Prasad, and M. M. Crawford, “Wavelet-domain multiview active learning for spatial-spectral hyperspectral

- image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4047–4059, 2016.
- [97] Z. Wang, B. Du, L. Zhang, L. Zhang, and X. Jia, "A novel semisupervised active-learning algorithm for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 6, pp. 3071–3083, June 2017.
- [98] S. Patra, K. Bhardwaj, and L. Bruzzone, "A spectral-spatial multicriteria active learning technique for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 12, pp. 5213–5227, Dec. 2017.
- [99] C. Liu, L. He, Z. Li, and J. Li, "Feature-Driven active learning for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 1, pp. 341–354, Jan. 2018.
- [100] X. Xu, J. Li, and S. Li, "Multiview intensity-based active learning for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 669–680, Feb. 2018.
- [101] C. Liu, J. Li, and L. He, "Super-pixel-Based semisupervised active learning for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 357–370, Jan. 2019.
- [102] Z. Zhang, E. Pasolli, and M. M. Crawford, "An adaptive multiview active learning approach for spectral-spatial classification of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 4, pp. 2557–2570, April 2020.
- [103] C. Mu, J. Liu, Y. Liu, and Y. Liu, "Hyperspectral image classification based on active learning and spectral-spatial feature fusion using spatial coordinates," *IEEE Access*, vol. 8, pp. 6768–6781, 03 January 2020.
- [104] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: an overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690–6709, 2019.
- [105] "What is deep learning and how does it work," <https://towardsdatascience.com/what-is-deep-learning-and-how-does-it-work-f7d02aa9d477>.
- [106] A. Subasi and S. Mian Qaisar, "The ensemble machine learning-based classification of motor imagery tasks in brain-computer interface," *Journal of Healthcare Engineering*, vol. 2021, Article ID 1970769, 12 pages, 2021.
- [107] B. Alsinglawi, O. Alshari, M. Alorjani et al., "An explainable machine learning framework for lung cancer hospital length of stay prediction," *Scientific Reports*, vol. 12, no. 1, pp. 1–10, 2022.
- [108] B. Alsinglawi, F. Alnajjar, O. Mubin, M. Novoa, O. Karajeh, and O. Darwish, "Benchmarking predictive models in electronic health records: sepsis length of stay prediction," in *Proceedings of the International Conference on Advanced Information Networking and Applications*, pp. 258–267, Springer, Caserta, Italy, 15 April 2020, https://link.springer.com/chapter/10.1007/978-3-030-44041-1_24.
- [109] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of skin disease using deep learning neural networks with MobileNet V2," *Sensors*, vol. 21, no. 8, p. 2852.
- [110] Z. Lin, Y. Chen, X. Zhao, and G. Wang, "Spectral-spatial classification of hyperspectral image using autoencoders," in *Proceedings of the 2013 9th International Conference on Information, Communications & Signal Processing*, IEEE, Tainan, 10 December 2013.
- [111] G. Liu, H. Bao, and B. Han, "A stacked autoencoder-based deep neural network for achieving gearbox fault diagnosis," *Advancements in Mathematical Methods for Pattern Recognition and its Applications*, vol. 2018, Article ID 5105709, 10 pages, 2018.
- [112] A. Tutorial, "A beginner's guide to autoencoders," <https://www.edureka.co/blog/autoencoders-tutorial/>.
- [113] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2094–2107, June 2014.
- [114] X. Ma, H. Wang, and J. Geng, "Spectral-spatial classification of hyperspectral image based on deep auto-encoder," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4073–4085, Sept. 2016.
- [115] P. Zhou, J. Han, G. Cheng, and B. Zhang, "Learning compact and discriminative stacked autoencoder for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4823–4833, July 2019.
- [116] H. Madani and K. McIsaac, "Distance transform-based spectral-spatial feature vector for hyperspectral image classification with stacked autoencoder," *Remote Sensing*, vol. 13, p. 1732, 2021.
- [117] "Layers of a convolutional neural network," <https://wiki.tum.de/display/Ifdv/Layers+of+a+Convolutional+Neural+Network>.
- [118] B. C. N. N. Architecture, "Explaining 5 layers of convolutional neural network," <https://www.upgrad.com/blog/basic-cnn-architecture/>.
- [119] "A survey of the recent architectures of deep convolutional neural networks," <https://arxiv.org/ftp/arxiv/papers/1901/1901.06032.pdf>.
- [120] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Date of Conference*, IEEE, Milan, Italy, 26 July 2015.
- [121] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
- [122] W. Zhao and S. Du, "Spectral-spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4544–4554, 2016.
- [123] J. Cao, Z. Chen, and B. Wang, "Deep convolutional networks with super-pixel segmentation for hyperspectral image classification," in *Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 10 July 2016.
- [124] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 844–853, Feb. 2017.
- [125] M. He, B. Li, and H. Chen, "Multi-scale 3d deep convolutional neural network for hyper-spectral image classification," in *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, Beijing, China, 17 September 2017.
- [126] X. Yang, Y. Ye, X. Li, R. Y. K. Lau, X. Zhang, and X. Huang, "Hyperspectral image classification with deep learning

- models,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5408–5423, 2018.
- [127] H. Zhang, Y. Li, Y. Jiang, P. Wang, Q. Shen, and C. Shen, “Hyperspectral classification based on lightweight 3-D-CNN with transfer learning,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5813–5828, 2019.
- [128] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, “HybridSN: exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 2, pp. 277–281, Feb. 2020.
- [129] W. Hu, H. Li, L. Pan, W. Li, R. Tao, and Q. Du, “Spatial-spectral feature extraction via deep con-vLSTM neural networks for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 6, pp. 4237–4250, June 2020.
- [130] M. E. Paoletti, J. M. Haut, S. K. Roy, and E. M. T. Hendrix, “Rotation equivariant convolutional neural networks for hyperspectral image classification,” *IEEE Access*, vol. 8, pp. 179575–179591, 2020.
- [131] X. Zhang, Y. Wang, N. Zhang et al., “Spectral-spatial three-dimensional convolutional neural network for hyperspectral image classification,” *IEEE Access*, vol. 8, pp. 127167–127180, 2020.
- [132] L. Mou, P. Ghamisi, and X. X. Zhu, “Deep recurrent neural networks for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3639–3655, July 2017.
- [133] “Types of RNN (recurrent neural network),” <https://iq.opengenus.org/types-of-rnn/>.
- [134] “5 types of LSTM recurrent neural networks and what to do with them,” <https://www.exactcorp.com/blog/Deep-Learning/5-types-of-lstm-recurrent-neural-networks-and-what-to-do-with-them>.
- [135] R. Hang, Q. Liu, D. Hong, and P. Ghamisi, “Cascaded recurrent neural networks for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5384–5394, 2019.
- [136] S. Hao, W. Wang, and M. Salzmann, “Geometry-aware deep recurrent neural networks for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing, Early Access*, vol. 59, pp. 1–13, 2020.
- [137] Y. Chen, X. Zhao, and X. Jia, “Spectral-spatial classification of hyperspectral data based on deep Belief network,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2381–2392, June 2015.
- [138] A. Mughees and L. Tao, “Multiple deep-belief-network-based spectral-spatial classification of hyperspectral images,” *Tsinghua Science and Technology*, vol. 24, no. 2, pp. 183–194, April 2019.
- [139] C. Chen, Y. Ma, and G. Ren, “Hyperspectral classification using deep Belief networks based on conjugate gradient update and pixel-centric spectral block features,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 4060–4069, 2020.
- [140] Y. Zhan, D. Hu, Y. Wang, and X. Yu, “Semisupervised hyperspectral image classification based on generative adversarial networks,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 212–216, Feb. 2018.
- [141] “The math behind GANs (generative adversarial networks),” <https://towardsdatascience.com/the-math-behind-gans-generative-adversarial-networks-3828f3469d9c>.
- [142] “Introduction to generative adversarial networks (GANs): types, and applications, and implementation,” <https://heartbeat.fritz.ai/introduction-to-generative-adversarial-networks-gans-35ef44f21193>.
- [143] L. Zhu, Y. Chen, P. Ghamisi, and J. A. Benediktsson, “Generative adversarial networks for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5046–5063, 2018.
- [144] H. Wang, C. Tao, J. Qi, H. Li, and Y. Tang, “Semi-supervised variational generative adversarial networks for hyperspectral image classification,” in *Proceedings of the IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, Yokohama, Japan, 28 July 2019.
- [145] C. Tao, H. Wang, J. Qi, and H. Li, “Semisupervised variational generative adversarial networks for hyperspectral image classification,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 914–927, Feb. 2020.
- [146] Z. Zhong, J. Li, D. A. Clausi, and A. Wong, “Generative adversarial networks and conditional random fields for hyperspectral image classification,” *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 3318–3329, July 2020.
- [147] H. Liang, W. Bao, and X. Shen, “Adaptive weighting feature fusion approach based on generative adversarial network for hyperspectral image classification,” *Remote Sensing*, vol. 13, p. 198, 2021.
- [148] Z. Li, X. Zhu, Z. Xin, F. Guo, X. Cui, and L. Wang, “Variational generative adversarial network with crossed spatial and spectral interactions for hyperspectral image classification,” *Remote Sensing*, vol. 13, p. 3131, 2021.
- [149] M. Tsiakmaki, G. Kostopoulos, S. Kotsiantis, and O. Ragos, “Transfer learning from deep neural networks for predicting student performance,” *Applied Sciences*, vol. 10, no. 6, p. 2145, 2020.
- [150] J. Lin, R. Ward, and Z. J. Wang, “Deep transfer learning for Hyperspectral Image classification,” in *Proceedings of the 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSp)*, IEEE, Yokohama, Japan, 29 August 2018.
- [151] C. Deng, Y. Xue, X. Liu, C. Li, and D. Tao, “Active transfer learning network: a unified deep joint spectral-spatial feature learning model for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 3, pp. 1741–1754, March 2019.
- [152] X. He, Y. Chen, and P. Ghamisi, “Heterogeneous transfer learning for hyperspectral image classification based on convolutional neural network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3246–3263, May 2020.
- [153] Y. Liu, L. Gao, C. Xiao, Y. Qu, K. Zheng, and A. Marinoni, “Hyperspectral image classification based on a shuffled group convolutional neural network with transfer learning,” *Remote Sensing*, vol. 12, p. 1780, 2020.
- [154] F. Xie, Q. Gao, C. Jin, and F. Zhao, “Hyperspectral image classification based on superpixel pooling convolutional neural network with transfer learning,” *Remote Sensing*, vol. 13, p. 930, 2021.
- [155] Y. Kumar, A. Koul, R. Singla, and M. F. Ijaz, “Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 1–28, 2022.
- [156] “Data-driven cervical cancer prediction model with outlier detection and over-sampling methods. MF Ijaz, M Attique, Y Son,” *Sensors*, vol. 20, no. 10, pp. 2809–76, 2020.

- [157] M. D. Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 3, pp. 542–546, May 2011.
- [158] J. Xia, J. Chanussot, P. Du, and X. He, "(Semi-) supervised probabilistic principal component analysis for hyperspectral remote sensing image classification," in *Proceedings of the 2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, IEEE, 4 June 2012.
- [159] Y. Ren, L. Liao, S. J. Maybank, Y. Zhang, and X. Liu, "Hyperspectral image spectral-spatial feature extraction via tensor principal component analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1431–1435, 2017.
- [160] S. Kutluk, K. Kayabol, and A. Akan, "Classification of hyperspectral images using mixture of probabilistic PCA models," in *Proceedings of the 2016 24th European Signal Processing Conference (EUSIPCO)*, IEEE, Budapest, Hungary, 29 August 2016.
- [161] X. Kang, X. Xiang, S. Li, and J. A. Benediktsson, "PCA-based edge-preserving features for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 7140–7151, Dec. 2017.
- [162] S. Chiang, C. Chang, and I. W. Ginsberg, "Unsupervised hyperspectral image analysis using independent component analysis," in *Proceedings of the IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No.00CH37120)*, IEEE, Honolulu, HI, USA, 24 July 2000.
- [163] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, "Independent component discriminant analysis for hyperspectral image classification," in *Proceedings of the 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, IEEE, Reykjavik, Iceland, 14–16 June 2010.
- [164] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, "Hyperspectral image classification with independent component discriminant analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 12, pp. 4865–4876, Dec. 2011.
- [165] Y. Chen, L. Xu, Y. Fang et al., "Unsupervised bayesian subpixel mapping of hyperspectral imagery based on band-weighted discrete spectral mixture model and Markov random field," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 1, pp. 162–166, Jan. 2021.
- [166] H. Chen, F. Miao, Y. Chen, Y. Xiong, and T. Chen, "A hyperspectral image classification method using multifeature vectors and optimized KELM," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 2781–2795, 2021.
- [167] A. Saboori, H. Ghassemian, and F. Razzazi, "Active multiple kernel Fredholm learning for hyperspectral images classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 2, pp. 356–360, Feb. 2021.
- [168] Y. Zhang, G. Cao, B. Wang, X. Li, P. Y. O. Amoako, and A. Shafique, "Dual sparse representation graph-based copropagation for semisupervised hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, 2021.
- [169] Z. Lei, Y. Zeng, P. Liu, and X. Su, "Active deep learning for hyperspectral image classification with uncertainty learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, 2021.
- [170] K. Y. Ma and C. I. Chang, "Iterative training sampling coupled with active learning for semisupervised spectral-spatial hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, 2021.
- [171] Q. Lu and L. Wei, "Multiscale superpixel-based active learning for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, 2021.
- [172] R. Hang, Z. Li, Q. Liu, P. Ghamisi, and S. S. Bhattacharyya, "Hyperspectral image classification with attention-aided CNNs," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 3, pp. 2281–2293, March 2021.
- [173] C. Yu, R. Han, M. Song, C. Liu, and C. I. Chang, "Feedback attention-based dense CNN for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, 2021.
- [174] R. Hang, F. Zhou, Q. Liu, and P. Ghamisi, "Classification of hyperspectral images via multitask generative adversarial networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 2, pp. 1424–1436, Feb. 2021.
- [175] W. Y. Wang, H. C. Li, Y. J. Deng, L. Y. Shao, X. Q. Lu, and Q. Du, "Generative adversarial capsule network with ConvLSTM for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 3, pp. 523–527, March 2021.
- [176] J. Wang, S. Guo, R. Huang, L. Li, X. Zhang, and L. Jiao, "dual-channel capsule generation adversarial network for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, 2021.
- [177] S. K. Roy, J. M. Haut, M. E. Paoletti, S. R. Dubey, and A. Plaza, "Generative adversarial minority oversampling for spectral-spatial hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, 2021.
- [178] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral images with a super-pixel-based discriminative sparse model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 8, pp. 4186–4201, 2015.
- [179] P. Ma, J. Ren, H. Zhao, G. Sun, P. Murray, and J. Zheng, "Multiscale 2-D singular spectrum analysis and principal component analysis for spatial-spectral noise-robust feature extraction and classification of hyperspectral images," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 1233–1245, 2021.
- [180] D. M. S. Arsa, H. R. Sanabila, M. F. Rachmadi, A. Gamal, and W. Jatmiko, "Improving principal component analysis performance for reducing spectral dimension in hyperspectral image classification," in *Proceedings of the 2018 International Workshop on Big Data and Information Security (IW BIS)*, pp. 123–128, IEEE, Jakarta, Indonesia, 12 May 2018.
- [181] M. Baisantriy and A. K. Sao, "Band selection using segmented PCA and component loadings for hyperspectral image classification," in *Proceedings of the IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 3812–3815, IEEE, Yokohama, Japan, 28 July 2019.
- [182] M. M. Hossain and M. A. Hossain, "Feature reduction and classification of hyperspectral image based on multiple kernel PCA and deep learning," in *Proceedings of the 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON)*, pp. 141–144, IEEE, Dhaka, Bangladesh, 29 November 2019.

- [183] D. Ruiz, B. Bacca, and E. Caicedo, "Hyperspectral images classification based on inception network and kernel PCA," *IEEE Latin America Transactions*, vol. 17, no. 12, pp. 1995–2004, December 2019.
- [184] J. Lin, L. Zhao, S. Li, R. Ward, and Z. J. Wang, "Active-learning-incorporated deep transfer learning for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 4048–4062, Nov. 2018.
- [185] J. Lin, L. Mou, X. X. Zhu, X. Ji, and Z. J. Wang, "Attention-aware pseudo-3-D convolutional neural network for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, 2021.
- [186] V. Neagoie and P. Diaconescu, "CNN hyperspectral image classification using training sample augmentation with generative adversarial networks," in *Proceedings of the 2020 13th International Conference on Communications (COMM)*, pp. 515–519, IEEE, Bucharest, Romania, 18 June 2020.
- [187] X. Zhang, X. Jiang, J. Jiang, Y. Zhang, X. Liu, and Z. Cai, "Spectral-spatial and superpixelwise PCA for unsupervised feature extraction of hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, 2021.
- [188] A. N. Abbasi and M. He, "Convolutional neural network with PCA and batch normalization for hyperspectral image classification," in *Proceedings of the IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 959–962, IEEE, Yokohama, Japan, 28 July 2019.
- [189] M. R. Haque and S. Z. Mishu, "Spectral-spatial feature extraction using PCA and multi-scale deep convolutional neural network for hyperspectral image classification," in *Proceedings of the 2019 22nd International Conference on Computer and Information Technology (ICCIT)*, pp. 1–6, IEEE, Dhaka, Bangladesh, 18 December 2019.
- [190] W. Sun, G. Yang, J. Peng, and Q. Du, "Lateral-slice sparse tensor robust principal component analysis for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 1, pp. 107–111, Jan. 2020.
- [191] A. N. Abbasi and M. He, "CNN with ICA-PCA-DCT joint preprocessing for hyperspectral image classification," in *Proceedings of the 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 595–600, IEEE, Lanzhou, China, 18 November 2019.
- [192] M. Baisantray, A. K. Sao, and D. P. Shukla, "Band selection using combined divergence–correlation index and sparse loadings representation for hyperspectral image classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5011–5026, 2020.
- [193] S. Jia, J. Hu, Y. Xie, L. Shen, X. Jia, and Q. Li, "Gabor cube selection based multitask joint sparse representation for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3174–3187, June 2016.
- [194] Y. Zhan, J. Qin, T. Huang et al., "Hyperspectral image classification based on generative adversarial networks with feature fusing and dynamic neighborhood voting mechanism," in *Proceedings of the IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 811–814, IEEE, Yokohama, Japan, 28 July 2019.
- [195] A. A. Joy and M. A. M. Hasan, "A hybrid approach of feature selection and feature extraction for hyperspectral image classification," in *Proceedings of the 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (ICAME2)*, pp. 1–4, IEEE, Rajshahi, Bangladesh, 11 July 2019.
- [196] U. A. M. E. Ali, M. A. Hossain, and M. R. Islam, "Analysis of PCA based feature extraction methods for classification of hyperspectral image," in *Proceedings of the 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET)*, pp. 1–6, Dhaka, Bangladesh, 23 December 2019.
- [197] Y. Zhan, K. Wu, W. Liu et al., "Semi-supervised classification of hyperspectral data based on generative adversarial networks and neighborhood majority voting," in *Proceedings of the IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5756–5759, IEEE, Valencia, Spain, 22 July 2018.
- [198] A. I. Champa, M. F. Rabbi, and N. Banik, "Improvement in hyperspectral image classification by using hybrid subspace detection technique," in *Proceedings of the 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp. 1–5, IEEE, Dhaka, Bangladesh, 24 December 2019.
- [199] H. Fu, G. Sun, J. Ren, A. Zhang, and X. Jia, "Fusion of PCA and segmented-PCA domain multiscale 2-D-SSA for effective spectral-spatial feature extraction and data classification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*.
- [200] G. Y. Chen, "Multiscale filter-based hyperspectral image classification with PCA and SVM," *Journal of Electrical Engineering*, vol. 72, no. 1, pp. 40–45, 2021.
- [201] Z. Xue, S. Zhou, and P. Zhao, "Active learning improved by neighborhoods and superpixels for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 3, pp. 469–473, March 2018.
- [202] K. Bhardwaj, A. Das, and S. Patra, "Spectral-spatial active learning with superpixel profile for classification of hyperspectral images," in *Proceedings of the 2020 6th International Conference on Signal Processing and Communication (ICSC)*, pp. 149–155, IEEE, Noida, India, 5 March 2020.
- [203] X. Cao, J. Yao, Z. Xu, and D. Meng, "Hyperspectral image classification with convolutional neural network and active learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 7, pp. 4604–4616, July 2020.
- [204] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Training capsnets via active learning for hyperspectral image classification," in *Proceedings of the IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, pp. 40–43, IEEE, Waikoloa, HI, USA, 26 September 2020.
- [205] J. M. Haut, M. E. Paoletti, J. Plaza, J. Li, and A. Plaza, "Active learning with convolutional neural networks for hyperspectral image classification using a new bayesian approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6440–6461, Nov. 2018.
- [206] X. Wang, K. Tan, and Y. Chen, "CapsNet and triple-GANs towards hyperspectral classification," in *Proceedings of the 2018 Fifth International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*, pp. 1–4, IEEE, Xi'an, China, 18 June 2018.
- [207] X. Wang, K. Tan, Q. Du, Y. Chen, and P. Du, "Caps-TripleGAN: GAN-assisted CapsNet for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 7232–7245, Sept. 2019.
- [208] O. Okwuashi and C. E. Ndehedehe, "Deep support vector machine for hyperspectral image classification," *Pattern Recognition*, vol. 103, Article ID 107298, 2020.
- [209] F. I. Alam, J. Zhou, A. W. Liew, X. Jia, J. Chanussot, and Y. Gao, "Conditional random field and deep feature learning

- for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 3, pp. 1612–1628, March 2019.
- [210] Y. Cao, J. Mei, W. Yuebin et al., “SLCRF: subspace learning with conditional random field for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, 2021.
- [211] Y. Liang, X. Zhao, A. J. X. Guo, and F. Zhu, “Hyperspectral image classification with deep metric learning and conditional random field,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 6, pp. 1042–1046, June 2020.
- [212] Y. Wang, J. Mei, L. Zhang et al., “Self-supervised feature learning with CRF embedding for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2628–2642, May 2019.
- [213] V. Andrejchenko, W. Liao, W. Philips, and P. Scheunders, “Decision fusion framework for hyperspectral image classification based on Markov and conditional random fields,” *Remote Sensing*, vol. 11, no. 6, p. 624, 2019.
- [214] J. V. Rissati, P. C. Molina, and C. S. Anjos, “Hyperspectral image classification using random forest and deep learning algorithms,” in *Proceedings of the 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)*, p. 132, IEEE, Santiago, Chile, 22 March 2020.
- [215] Y. Zhang, G. Cao, X. Li, and B. Wang, “Cascaded random forest for hyperspectral image classification,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 4, pp. 1082–1094, April 2018.
- [216] Y. Zhang, G. Cao, X. Li, B. Wang, and P. Fu, “Active semi-supervised random forest for hyperspectral image classification,” *Remote Sensing*, vol. 11, no. 24, p. 2974, 2019.
- [217] T. Li, J. Leng, and L. D. C. N. R. Kong, “Deep cube CNN with random forest for hyperspectral image classification,” *Multimedia Tools and Applications*, vol. 78, pp. 3411–3433, 2019.
- [218] A. Wang, Y. Wang, and Y. Chen, “Hyperspectral image classification based on convolutional neural network and random forest,” *Remote Sensing Letters*, vol. 10, no. 11, pp. 1086–1094.
- [219] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, “Graph convolutional networks for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, 2021.
- [220] Y. Ding, Y. Guo, Y. Chong, S. Pan, and J. Feng, “Global consistent graph convolutional network for hyperspectral image classification,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–16, Article ID 5501516, 2021.
- [221] X. He, Y. Chen, and P. Ghamisi, “Dual graph convolutional network for hyperspectral image classification with limited training samples,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, 2021.
- [222] J. Bai, B. Ding, Z. Xiao, L. Jiao, H. Chen, and A. C. Regan, “Hyperspectral image classification based on deep attention graph convolutional network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, 2021.
- [223] Q. Liu, L. Xiao, J. Yang, and Z. Wei, “CNN-enhanced graph convolutional network with pixel- and superpixel-level feature fusion for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, 2021.
- [224] A. Signoroni, M. Savardi, A. Baronio, and S. Benini, “Deep learning meets hyperspectral image analysis: a multidisciplinary review,” *Journal of Imaging*, vol. 5, no. 5, p. 52, 2019.

Research Article

English Teaching Quality Monitoring and Multidimensional Analysis Based on the Internet of Things and Deep Learning Model

Juan Song 

Chongqing Metropolitan College of Science and Technology, Yongchuan, Chongqing 402167, China

Correspondence should be addressed to Juan Song; songjuan@cqcst.edu.cn

Received 23 February 2022; Revised 23 March 2022; Accepted 25 March 2022; Published 28 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Juan 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.

With the development of the times, English as the universal language in the world has been highly valued by the society and schools, and English skills have become a basic skill in the society. The school is actively developing, and in the process of reform and development, the monitoring of teaching quality is essential. Teaching quality is a complex and vague concept. The establishment of a teaching quality monitoring system helps to ensure the quality of personnel training and improve the level of education and teaching, and the quality of classroom teaching is the core content of education quality. Teaching quality monitoring is the management process of various measures and actions taken to ensure the continuous improvement of students' learning quality and to achieve certain quality standards by systematically supervising and controlling various factors affecting the teaching quality in the teaching process. The research results of the article show that (1) under the traditional teaching mode, the average grades of the three groups were 74, 72, 67, 62, and 62, respectively. The score of the oral test module is relatively low, the highest score is only 63 points, the overall score shows a low level, the academic achievement is hovering on the edge of passing, and the students' English learning situation is poor. Under the new classroom quality monitoring mode, the average scores of the three groups were 96, 92, 90, 86, and 84, respectively. Compared with the ordinary teaching mode, the scores of the five detection modules were greatly improved, and the average score of the listening module was improved. The teaching contents include 22 listening modules, 20 reading modules, 23 cloze modules, 24 translation modules, and 22 speaking modules. Overall, students' English learning level has been greatly improved. (2) Generally speaking, the overall reliability coefficient of the sample data with full English teaching content is mostly kept in the range of 0.70–0.95, and only a few parts show a low situation, which also shows the overall situation of the diversity of teaching content. The overall reliability coefficient of the sample data of good teaching methods shows a relatively high situation, the reliability coefficient of the improvement of learning interest can reach 0.93, the reliability coefficient of the understanding of the learning content can reach the highest 0.96, the problem analysis ability can reach the highest 0.97, and the innovation ability can reach 0.96. The improved reliability coefficient can reach up to 0.98, which shows the authenticity and validity of the experimental data. (3) The detection result of the new classroom quality monitoring teaching mode is the highest among several models, the accuracy rate can reach 96.42%, the recall rate can reach 97.21%, and the F1 value can reach 97.46%, indicating that the new classroom quality monitoring teaching mode is effective. Teaching performance is the highest. According to the ROC curve values of the four models, we can also conclude that the ROC value of the new classroom quality monitoring teaching has been maintained at 0.98 without major twists and turns. Whether it is in the test set or the training set, the detection results of the new classroom quality monitoring are still the highest, the accuracy rate can reach 94.42%, the recall rate can reach 94.78%, and the F1 value can reach 94.49%. After the training set runs, except the performance of the traditional teaching mode increases, the performance of the other 3 models decreases.

1. Introduction

In order to promote the development of the global economy, the communication between all walks of life in the country is inevitable. As the carrier of communication

between people, the free switching and transmission of languages in different countries are particularly important. The more complex grammar problems make the English learning process more difficult. In recent years, English has made continuous progress and innovation in education in

my country. The school has also changed the original teaching mode, actively improved the teaching ability of teachers, strived to cultivate students' innovative ability, and constantly developed new teaching methods and improved the quality of English teaching. In teaching, English teachers should choose appropriate teaching methods, which play a key role in improving teachers' teaching ability and improving students' English performance. According to the situation of different students, we should teach students according to their aptitude and improve the quality of teaching. Literature [1] studied the theory of deep learning and proposed a blended learning strategy to encourage students to achieve deep learning. Reference [2] revolved around deep learning model-based IICS anomaly detection techniques using information collected from TCP/IP packets for learning and validation. In Reference [3], a computer framework facilitates the inference of distributed deep learning models and is executed cooperatively by devices in a distributed computing hierarchy. Reference [4] showed that vernacular teaching reduces the value of the reflective practitioner spirit among teachers and thus hinders the initiative of CPD among university teachers. Literature [5] studied the application of data mining technology in teaching system management, discussed the current teaching quality evaluation system in colleges and universities, and proposed that the embedded system in the Internet of Things can transmit the data from the sensor layer to the network layer, so as to realize the integration between the sensor and the Internet. Quality management is the core task of teaching, and the establishment of monitoring system is an important content of quality management [6]. The quality of learning and teaching depends on many factors, the most important being learner self-assessment and formal assessment or testing, monitoring learner success and achievement, teacher development, and teacher evaluation[7]. Reference [8] proposed a deep learning-based IoT real-time health monitoring system using cross-testing to extensively evaluate the performance of the proposed system. Reference [9] introduced two different machine learning-based algorithms for photovoltaic (PV) array fault monitoring and classification. The study demonstrated the capability and feasibility of the method. Reference [10] illustrated a brand-new listing and oral course model based on the Tell Me More teaching system based on a blended English teaching model. Literature [11] studied and analyzed the influencing factors of college English classroom teaching quality evaluation and formed a relatively complete and suitable college English classroom in China. Reference [12] illustrated some potential and noteworthy issues and advantages of e-learning models applied to English teaching. Reference [13] demonstrated the feasibility of building IoT applications powered by effective, efficient, and reliable deep learning models. Reference [14] emphasized the importance of the English teaching quality monitoring system. Literature [15] proposed three aspects: teachers' professional ethics and habits, giving full play to teachers' leading role, and improving students' performance evaluation system and English teaching.

2. English Teaching Quality Monitoring and Analysis

2.1. Status Quo of English Teaching Quality. There are many factors in the process of English teaching. First of all, the purpose of English teaching must be clear, so that students can gain some perceptual knowledge of English, stimulate their interest in English learning and cultivate their ability, and encourage students to speak English boldly, so as to lay a solid foundation, to develop good study habits to lay the foundation for further study of English. In the process of English teaching, we cannot rely solely on teachers' simple explanations and students' listening methods. We should change the traditional teaching methods and have appropriate oral interaction with students. Teachers should update English teaching concepts appropriately in teaching, further study the teaching concepts in the English curriculum standards, highlight the practicality of language, and should use less mechanical memory and more flexible use in teaching. Taking advantage of the Internet, it is possible to learn English universally anytime, anywhere. An English teaching quality monitoring system based on the Internet of Things and deep learning models is established, and a scientific and standardized comprehensive teaching quality management system is built.

2.2. Teaching Quality Monitoring System. The teaching quality assurance system in colleges and universities is to use the concepts and methods of system theory to organize the functions of various stages and links of quality management in order to achieve the training objectives and to implement institutionalized, structured, and continuous monitoring of personnel training activities. The teaching process is evaluated and diagnosed, and a stable and effective quality management system is formed with clear tasks, responsibilities, and authorities that coordinate and promote each other to ensure and improve the quality of teaching. Specifically implemented in the implementation process of college management, it mainly includes the teaching objective system, the curriculum system of colleges and universities, the teacher system of colleges and universities, and the evaluation system. To establish a perfect teaching quality assurance system in colleges and universities, it is necessary to ensure the effective and harmonious development of these subsystems. To ensure that the objects you monitor can reach the standard in every teaching link, only when every small link in the front meets the standard and lays a good foundation, you can reap the full benefits in the final major link. In this way, under the complete set of scientific teaching quality monitoring system, the teaching process can be improved towards the established goals and its shortcomings can be improved to form high-efficiency and high-quality teaching. The evaluation indicators for constructing the teaching quality of efficient teachers are shown in Table 1.

2.3. The Role of the Teaching Monitoring System in Teaching. Through the cloud education and teaching quality evaluation and monitoring system, scientific analysis and diagnosis

TABLE 1: Teaching quality evaluation indicators.

First-level indicator	Secondary indicators
Teaching content	The teaching objectives are clear and meet the teaching content requirements. Teachers have a thorough understanding of basic theories, concepts, key points, and difficult points [16].
	The content is substantial, the arrangement is reasonable, and the introduction of new achievements is emphasized [17].
Teaching skills	The layout of the blackboard writing is reasonable, the levels are clear, and the content is clear. The language is fluent, accurate, and clear.
	Teachers adopt a variety of teaching methods, teach students in accordance with their aptitude, and focus on cultivating students' ability to analyze and solve problems [18].
Teaching attitude	The teaching work is enthusiastic, full of energy, serious, and devoted. Abide by the teaching discipline, get to and from class on time.
	Teachers are rigorous in their studies, strict in teaching, and well-mannered.
Teacher's knowledge level and research ability	The knowledge is extensive, and the teaching can absorb the theory and knowledge of multiple courses.
	Have high scientific research ability. Have a certain scientific research ability.
Appraisal and performance	The assessment indicators are scientific and adopt a combination of diagnostic, procedural, and summative methods [19].
	Carefully correct students' experimental reports, with assessment analysis and summary. The classroom atmosphere is active and students are more motivated to learn.
	Students can operate experiments independently and correctly and have a solid grasp of basic knowledge [20].

of the effects and existing deficiencies of all aspects of teaching are carried out to provide reference for teaching decision-making or improvement [21]. It reflects the teaching effect of teachers and the learning effect of students, so as to encourage them to further improve their teaching ability and learning methods. For teachers, it is difficult for teachers to find their own shortcomings in the teaching process. Teaching monitoring can solve these problems very well. It can not only enable teachers to find their own problems in time, correct them, and create new teaching methods, but also form efficient and high-quality English classroom teaching. It is also possible to continuously tap the potential of teachers, understand their own advantages, and carry them forward. For students, complete teaching quality standards can help students better complete teaching tasks, guide students in the direction of learning, make each student master their own learning methods, and formulate different learning foundations for each student. The goal of learning is to enable students to study with high quality and high efficiency.

3. Quality Monitoring of English Teaching Based on Internet of Things Technology

3.1. *Establishment of Monitoring Model.* The English teaching quality monitoring adopts the maximum and minimum value method in normalization, which can well preserve the original meaning of the data. The calculation formula is as follows:

$$X = \frac{I - I_{\min}}{I_{\max} - I_{\min}}. \quad (1)$$

Among them, X is the normalized score, I_{\min} is the minimum teaching quality score, I_{\max} is the maximum score, and I is the unprocessed score.

English learning error formula is as follows [22]:

$$E = \frac{\sum_{k=1}^p \sum_{j=1}^l (y_j^k - o_j^k)}{2}, \quad (2)$$

$$\text{fitness} = \frac{1}{E}.$$

The formula for calculating the probability of selection is as follows [23]:

$$z(a_j) = \frac{f(a_j)}{\sum_{j=1}^d f(a_j)}. \quad (3)$$

English teaching quality monitoring is as follows:

$$q(a_k) = \sum_{j=1}^k z(a_j). \quad (4)$$

Sample of classroom teaching quality evaluation is as follows [24]:

$$E = \frac{1}{n} \sum_{j=1}^n \left| \sum_{k=1}^m (B_{jk} - A_{jk})^2 \right|. \quad (5)$$

Normalized processing is as follows:

$$X^*(k) = \frac{x_k - x_{k\min}}{x_{k\max} - x_{k\min}}. \quad (6)$$

The original English teaching quality evaluation data are standardized, and the calculation is published as follows:

$$x'_{ij} = \frac{(x_{ij} - \bar{x})}{S_j}. \quad (7)$$

Normalized value is as follows [25]:

$$Z_{ij} = x'_{ij} + A. \quad (8)$$

The same quantification of teaching quality evaluation indicators is as follows:

$$p_{ij} = \frac{Z_{ij}}{\sum_{i=1}^m Z_{ij}}, \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n). \quad (9)$$

The index entropy value E_j is calculated as follows:

$$E_j = -k \sum_{i=1}^m p_{ij} \ln(p_{ij}), \quad (10)$$

in

$$k = \frac{1}{\ln(n)}, \quad E_j \geq 0. \quad (11)$$

The coefficient of variance G_j is calculated as follows:

$$G_j = 1 - E_j. \quad (12)$$

Indicator weight w_j is calculated as follows:

$$w_j = \frac{G_j}{\sum_{j=1}^n G_j}. \quad (13)$$

The teaching quality of the sample F_i is calculated as follows:

$$F_i = \sum_{j=1}^n w_j p_{ij}. \quad (14)$$

3.2. Establishment of English Teaching Quality Model. Calculating the normalization of English teaching quality indicators,

$$\bar{u}_{ij} = \frac{u_{ij}}{\sum_{k=1}^n u_{kj}}. \quad (15)$$

The canonical average is found as follows:

$$\hat{w}_i = \frac{1}{n} \sum_{j=1}^n u_{ij}. \quad (16)$$

English teaching evaluation index is as follows:

$$CI = \frac{\lambda_{\max} - n}{n - 1}. \quad (17)$$

Final evaluation index is as follows:

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n \frac{(U\hat{W})}{\hat{W}_i}. \quad (18)$$

4. Simulation Experiments

4.1. Comparative Experiment. In order to make the experimental data more convincing, we selected 3 teachers with the same teaching age to teach the students in the 3 groups, respectively, and the teaching duration was one academic year. The experiment compares the student achievement of the traditional teaching mode with the student achievement of the new classroom quality monitoring mode and observes the superiority of the new classroom quality monitoring mode in English education. Before the start of the experiment, in order to ensure the objectivity of the experimental data, the three groups were tested separately, including the written part (listening, reading, gestalt, translation, and composition) and the oral part. The written part was graded by two teachers, and the oral part was taken. Generally, the average score of two teachers is taken. The specific experimental data are shown in Table 2 and Figure 1.

From the data in Table 2, we can conclude that under the traditional teaching mode, the average scores of the experimental group, the control group, and the standard group are 74, 72, 67, 62, and 62, respectively, and there is no significant difference in the scores of the three groups. Among them, the score of the experimental group is the highest among the three groups, and the listening modules of the three groups are the highest among all the detection modules. The listening score of the experimental group can reach 74 points, and the listening score of the control group is 74 points; the listening score of the standard group is 72 points. The score of the oral test module is relatively low, the highest score is only 63 points, the overall score shows a low level, the academic achievement is hovering on the edge of passing, and the students' English learning situation is poor.

According to the data in Table 3 and Figure 2, we can conclude that under the new classroom quality monitoring mode, the average scores of the experimental group, the control group, and the standard group are 96, 92, 90, 86, and 84, respectively. The speaking module is still the highest among the testing modules. The oral score of the experimental group is 98, the score of the control group is 96, and the score of the standard group is 94. Compared with the traditional teaching model, the scores of the five testing modules are all there was a big improvement, with a 22 average improvement in listening modules, a 20 average improvement in reading modules, a 23 average improvement in cloze modules, a 24 average improvement in translation and writing modules, and a 24 average improvement in speaking modules. The score is 22. Generally speaking, the English learning level of the students has been greatly improved. The experimental results also show that the new classroom quality monitoring mode can improve the teaching quality.

4.2. Simulation Experiment. The experiment combines the results of teachers' English teaching evaluation to find out the factors that affect English teaching. The experiment adopts the

TABLE 2: Results statistics of traditional teaching mode.

Group	Hearing	Hearing	Cloze	Translation and writing	Oral
Test group	76	74	68	64	63
Control group	74	72	68	62	62
Standard group	72	70	65	60	61
Average score	74	72	67	62	62

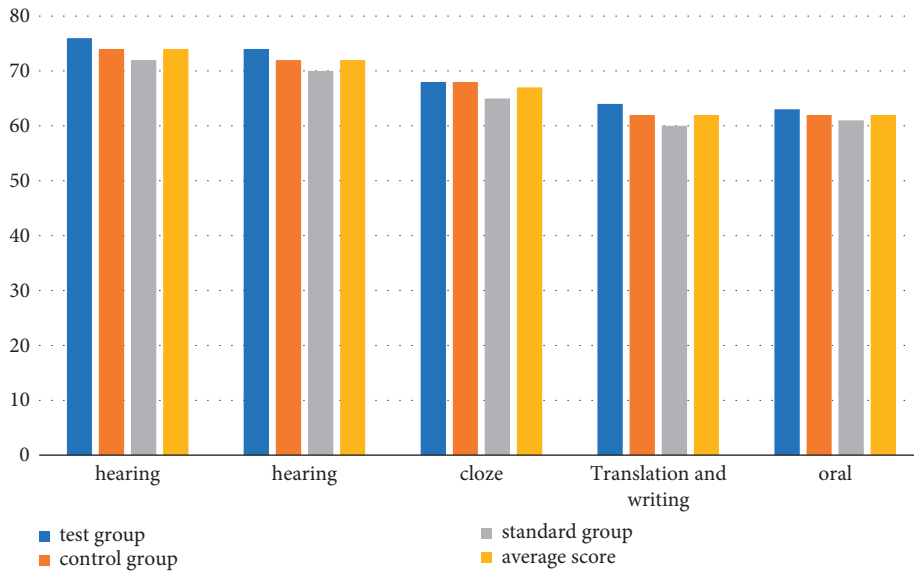


FIGURE 1: Statistics of the traditional teaching mode.

TABLE 3: New classroom quality monitoring mode.

Group	Hearing	Hearing	Cloze	Translation and writing	Oral
Test group	98	94	92	88	86
Control group	96	92	90	86	84
Standard group	94	90	88	84	82
Average score	96	92	90	86	84

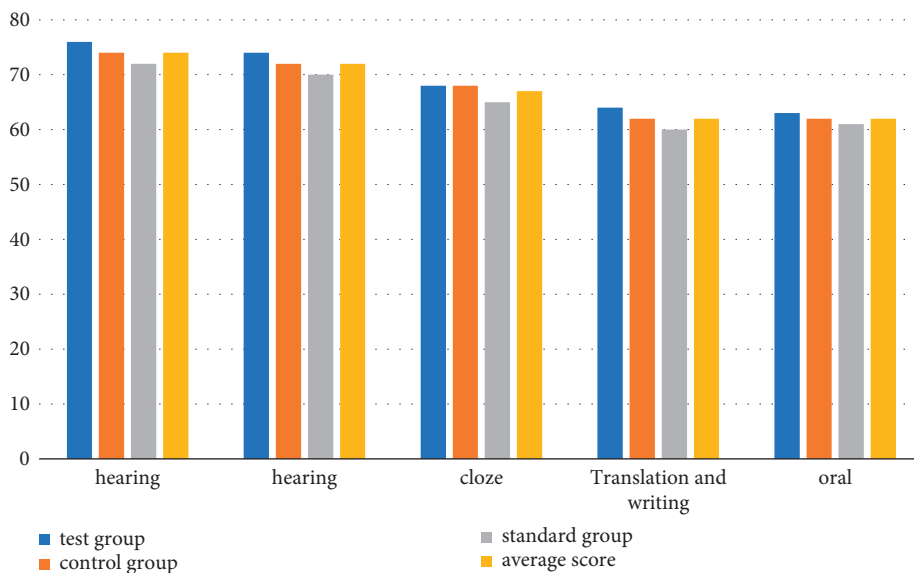


FIGURE 2: Results statistics of the new classroom quality monitoring model.

TABLE 4: Sample data of teaching content.

Sample number	Concept theory is accurate	English teaching content is full	Focus on practice	Depth of English knowledge
1	0.95	0.96	0.91	0.85
2	0.72	0.73	0.81	0.59
3	0.92	0.79	0.65	0.71
4	0.68	0.81	0.86	0.92
5	0.75	0.72	0.71	0.83
6	0.70	0.63	0.74	0.67
7	0.82	0.87	0.99	0.90
8	0.91	0.92	0.68	0.95
9	0.88	0.62	0.61	0.65
10	0.70	0.62	0.66	0.65

method of questionnaire survey. According to the collected questionnaire results, the results are divided into two indicators. The first-level indicators can be subdivided into five. Grades and secondary indicators can be subdivided into 18 grades, the developed questionnaires are scored by students, and the results are collected to obtain experimental data. The experiment is scored from two aspects: teaching content and teaching method. The experimental results are as follows.

4.2.1. Teaching Content. According to the data in Table 4 and Figure 3, we can conclude that the conceptual theoretical accuracy reliability coefficient of sample No. 1 has reached 0.95, and the other reliability coefficients of No. 1 are also maintained above 0.85, indicating the authenticity of the experimental data. In general, the overall reliability coefficient of the sample data with full English teaching content is mostly kept in the range of 0.70–0.95, and only a few parts show a low situation, which also shows that the overall situation of the diversity of teaching content is good. The reliability coefficient of practice is generally in a high state, and the reliability coefficient of deep understanding of English knowledge can reach a maximum of 0.95. In general, the test results of teachers' teaching content are in a good state, and the diversity of teachers' teaching content is good.

4.2.2. Teaching Methods. According to the data in Table 5 and Figure 4, we can conclude that the overall reliability coefficient of the teaching method sample data shows a high situation, the reliability coefficient of the improvement of problem analysis ability can reach 0.93, and the reliability coefficient of the understanding of the learning content is the highest. Reaching 0.96, the analysis ability can be improved up to 0.97, and the reliability coefficient of innovation ability can be up to 0.98, which shows the authenticity and validity of the experimental data. Among them, the reliability coefficient of sample data 1 shows a high state.

4.3. Model Performance Check. In the experiment, the model proposed in the article and other teaching models are run in different dimensions to test the superiority of the model. The experimental method is to run the four models in two test sets in turn. Among them, the training set is the sample set which is set aside during the model training process, which can be used to adjust the hyperparameters of the model and

evaluate the ability of the model. However, the test data set is different. Although it is the same sample set which is set aside during the model training process, it is used to evaluate the performance of the final model, helping to compare multiple final models and make choices. The ROC curve combines the sensitivity and specificity in a graphical way, which can accurately reflect the relationship between the specificity and sensitivity of the analytical method. It is a comprehensive representative of the test accuracy and can represent the performance of each model in the article. The specific experimental data are shown in Tables 6 and 7.

According to the data in Table 6 and Figure 5, we can conclude that the detection result of the new classroom quality monitoring teaching mode proposed by the article is the highest among several models, the accuracy rate can reach 96.42%, the recall rate can reach 97.21%, and the F1 value can reach 97.46%, indicating that the teaching performance of the new classroom quality monitoring teaching mode is the highest. Among them, the experimental results of the traditional teaching mode show a lower state, which indicates that the students' English learning effect is not ideal under the traditional teaching mode. The detection values of BP neural network teaching and adaptive teaching mode remain in the middle of the two. According to the ROC curve values of the four models, we can also conclude that the ROC value of the new classroom quality monitoring teaching has been kept at 0.98 without major twists and turns. The ROC curve of the BP neural network teaching is more tortuous, and the ROC value is relatively unstable. The description is less accurate.

According to the data in Table 7 and Figure 6, we can conclude that, whether in the test set or the training set, the detection results of the new classroom quality monitoring are still the highest, the accuracy rate can reach 94.42%, the recall rate can reach 94.78%, and the F1 value can reach 94.49%. After the training set runs, the detection effect of the model decreases to a certain extent, and the detection result of the traditional method improves to a certain extent due to the long use time. According to the ROC curves of the four models, we can also see that the ROC value of the new classroom quality monitoring has been stable at 0.98 without major twists and turns. The ROC curves of the other three models are more tortuous, and the ROC values are lower. It also shows that the recognition accuracy of the new classroom quality monitoring model is the highest.

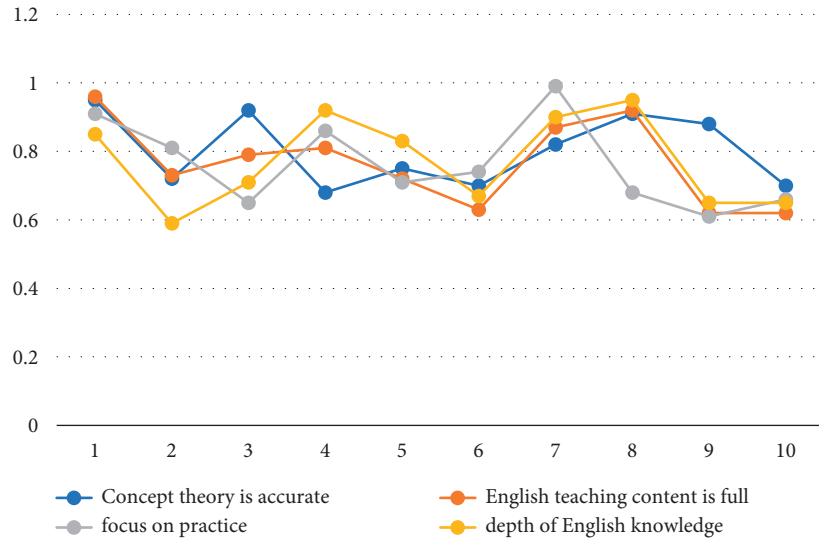


FIGURE 3: Sample results of teaching content.

TABLE 5: Sample data of teaching methods.

Sample number	Increased interest in learning	Understand the learning content	Understand the learning content	Improved ability to innovate
1	0.83	0.92	0.94	0.87
2	0.86	0.84	0.81	0.85
3	0.87	0.62	0.83	0.82
4	0.93	0.82	0.91	0.87
5	0.85	0.86	0.73	0.68
6	0.79	0.63	0.87	0.89
7	0.93	0.92	0.97	0.92
8	0.83	0.96	0.97	0.98
9	0.79	0.89	0.68	0.79
10	0.82	0.65	0.68	0.76

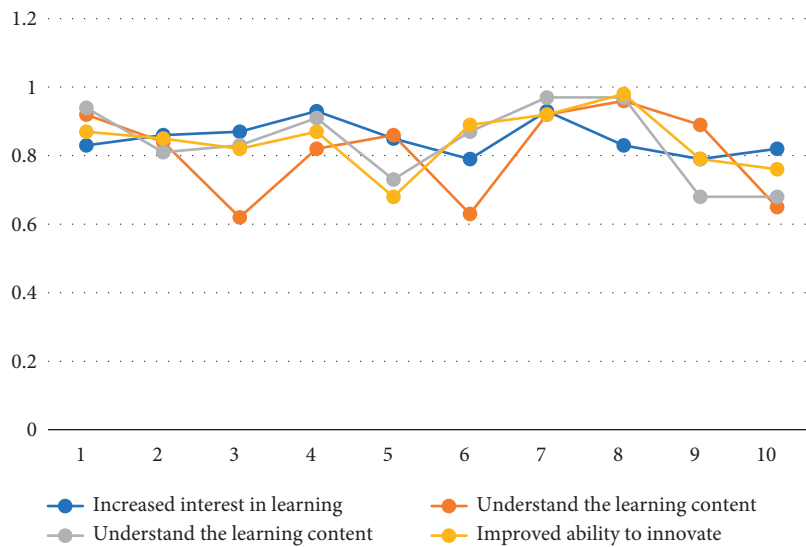


FIGURE 4: Teaching method sample results.

TABLE 6: Test set experimental results.

Model	Accuracy (%)	Recall (%)	F1 score (%)
New classroom quality monitoring	96.42	97.21	97.46
BP neural network teaching	92.41	93.12	93.24
Adaptive teaching mode	88.24	87.14	84.23
Traditional teaching mode	79.21	80.24	81.23

TABLE 7: Experimental results of training set.

Model	Accuracy (%)	Recall (%)	F1 score (%)
New classroom quality monitoring	94.42	94.78	94.49
BP neural network teaching	90.23	90.92	91.24
Adaptive teaching mode	85.46	85.89	86.19
Traditional teaching mode	82.46	83.10	83.78

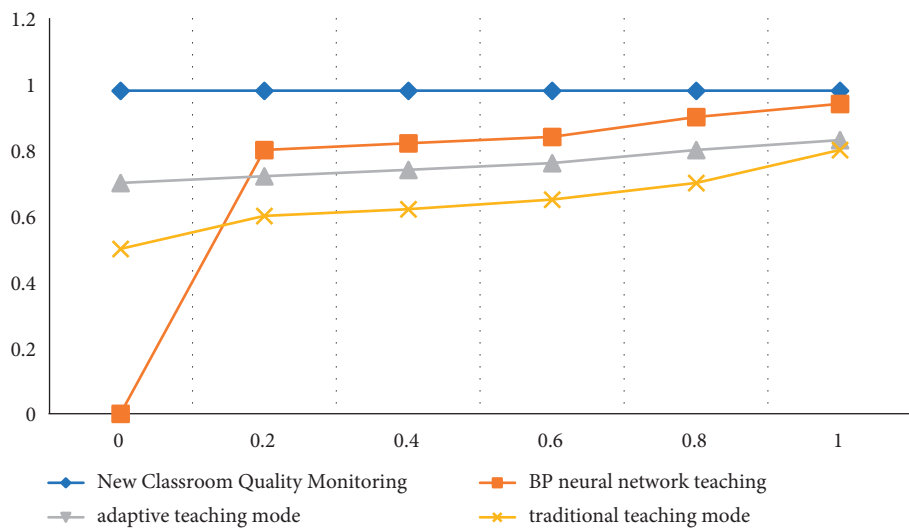


FIGURE 5: ROC curve on the test set.

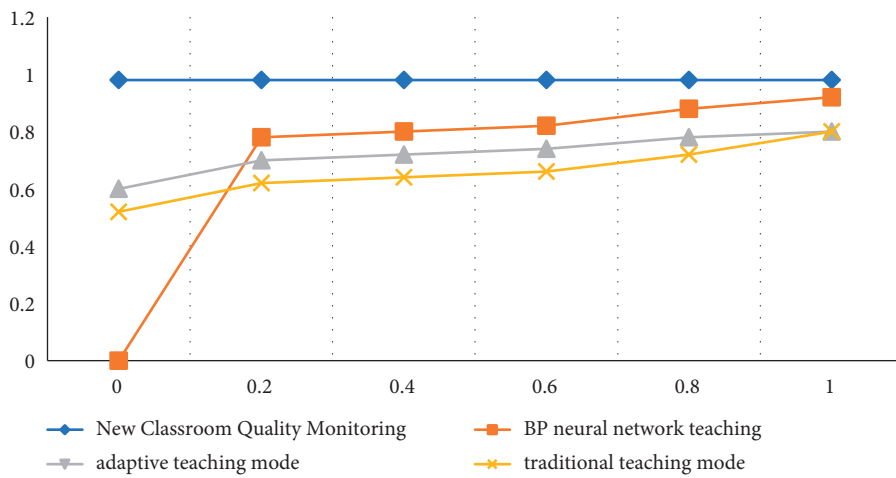


FIGURE 6: ROC curve on the training set.

5. Conclusion

English has received extensive attention from society and schools as the world's universal language. English has been regarded as a basic skill to enter the society, but some universities still have the problem of English teaching defects, because English is not our mother tongue, the particularity of English teaching, English teaching has not received enough attention, schools lack excellent teacher resources, and students cannot enjoy high-quality education. Since English involves a lot of grammar problems, there are many problems in all English teaching process. All students should gradually cultivate their English reading habits. They can listen to more English songs and watch English movies, which is convenient and better to improve the effect of classroom teaching. Although the English teaching quality monitoring model in this paper can monitor the English learning situation, it needs to be further improved in terms of teaching evaluation. We believe that the establishment of a scientific college English teaching quality monitoring system will not only improve the quality of college English teaching but also promote the rapid integration of Chinese college education with international college education and can also avoid language barriers.

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 author declares that there are no conflicts of interest regarding this work.

References

- [1] Y. Wang, "Towards a blended learning model based on deep learning theory: practice and research," *Distance Education in China*, vol. 10, no. 3, pp. 21–29, 2013.
- [2] M. Al-Hawawreh, N. Moustafa, and E. Sitnikova, "Identification of malicious activities in industrial internet of things based on deep learning models," *Journal of Information Security and Applications*, vol. 41, pp. 1–11, 2018.
- [3] M. H. Cheng, Q. Sun, and C. H. Tu, "An adaptive computation framework of distributed deep learning models for internet-of-things applications," in *Proceedings of the 2018 IEEE 24th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)*, vol. 12, no. 3, pp. 11–17, IEEE, Hakodate, Japan, August 2018.
- [4] K. Dou, "Establishment of monitoring and evaluation system of business English teaching in higher vocational colleges," *Vocational Technology*, vol. 08, no. 14, pp. 1–8, 2017.
- [5] C. H. Tu, Q. H. Sun, and M. H. Cheng, "On designing the adaptive computation framework of distributed deep learning models for Internet-of-Things applications," *The Journal of Supercomputing*, vol. 77, no. 1, pp. 17–25, 2021.
- [6] H. Wang, "Teaching quality monitoring and evaluation using 6G internet of things communication and data mining," *International Journal of System Assurance Engineering and Management*, no. 3, pp. 1–8, 2002.
- [7] G. Kavaliauskienė, "Quality assessment in teaching English for specific purposes," *English for Specific Purposes World*, vol. 12, no. 5, pp. 23–35, 2004.
- [8] X. Wu, C. Liu, and L. Wang, "Internet of things-enabled real-time health monitoring system using deep learning," *Neural Computing & Applications*, vol. 13, no. 8, pp. 1–12, 2021.
- [9] A. Mellit, O. Herrak, and C. R. Casas, "A machine learning and internet of things-based online fault diagnosis method for photovoltaic arrays," *Sustainability*, vol. 13, no. 4, pp. 14–23, 2021.
- [10] College of Mechanical Engineering and Automation, "Henan polytechnic institute nanyang, China department of computer science and engineering, jaypee university of information technology solan India. Research on transformer vibration monitoring and diagnosis based on internet of things," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 677–688, 2021.
- [11] L. Panpan, "Study on monitoring index system of college English classroom teaching quality based on BOPPPS model," in *Proceedings of the 9th International Conference on Education and Social Science (ICESS 2019)*, pp. 873–877, Sheenyang, Liaoning, China, Francis Academic Press, March 29, 2019.
- [12] A. Shaparenko and A. Golikova, "Problems and advantages of applying the e-learning model to the teaching of English," *University of Surrey*, vol. 10, no. 7, pp. 25–36, 2013.
- [13] S. Yao, Y. Zhao, A. Zhang et al., "Deep learning for the internet of things," *Computer*, vol. 51, no. 5, pp. 32–41, 2018.
- [14] H. Yin, "Research on the quality of college English teaching design," *Based on BOPPPS Model*, vol. 08, no. 12, pp. 56–65, 2018.
- [15] J. F. Zang, "On how to improve teaching quality monitoring abilities," *Journal of Shandong Institute of Commerce and Technology*, vol. 3, no. 10, pp. 55–63, 2014.
- [16] F. Wang and L. Zhao, "Research and practice of teaching quality monitoring system in applied undergraduate colleges and departments," *China Business Circle*, vol. 14, no. 8, pp. 130–133, 2009.
- [17] X. Li and S. Gao, "Research on the quality monitoring index system of college English classroom teaching," *China Adult Education*, no. 12, pp. 151–152, 2010.
- [18] E. Jana, V. MaryEllen, and J. S. Deborah, *Making Content Comprehensible for English Learners: the SIOP mode*, Pearson/Allyn&Bacon, Boston, MA, USA, 2008.
- [19] D. Jonassen and P. Hanning, "Mental models: knowledge in the head and knowledge in the world," *Educational Knowledge*, vol. 39, no. 5/6, pp. 37–41, 1999.
- [20] E. Jana and Powers Kristin, "School reform and standard-based education: A model for English language learners," *The Journal of Educational Research*, vol. 99, no. 4, pp. 23–32, 2006.
- [21] H. R. Abadiano and J. Turner, "Sheltered Instruction: an empowerment framework for English language learners," *The New England Reading Association Journal*, vol. 38, no. 3, pp. 36–41, 2002.
- [22] M. An, "The reform of college English teaching in Beijing," *Foreign Language Teaching and Research*, vol. 27, no. 1, pp. 16–18, 1999.

- [23] W. Sun, *University Classroom: Don't Rest Easy—Thinking Caused by the Large-Scale Undergraduate Classroom Teaching Inspection in Beijing*, China Higher Education, Suzhou, China, 1999.
- [24] (English) Wilkins, *The Principle of Foreign Language Learning and Teaching*, International Culture Publishing Company, Beijing, China, 1987.
- [25] C. Yang and M. Liu, "College English teachers' information technology ability and cultivation," *Foreign Languages*, vol. 26, no. 4, pp. 23–34, 2006.

Research Article

Automatic Gray Image Coloring Method Based on Convolutional Network

Jiayi Fan , Wentao Xie, and Tiantian Ge

Suzhou Institute of Technology, Jiangsu University of Science and Technology, Zhangjiagang, Jiangsu 215600, China

Correspondence should be addressed to Jiayi Fan; 999620150037@just.edu.cn

Received 7 March 2022; Revised 30 March 2022; Accepted 4 April 2022; Published 26 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Jiayi 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.

Image coloring is a time-consuming and laborious work. For a work, color collocation is an important factor to determine its quality. Therefore, automatic image coloring is a topic with great research significance and application value. With the development of computer hardware, deep learning technology has achieved satisfactory results in the field of automatic coloring. According to the source of color information, this paper can divide automatic coloring methods into three types: image coloring based on prior knowledge, image coloring based on reference pictures, and interactive coloring. The coloring method can meet the needs of most users, but there are disadvantages such as users cannot get the multiple objects in a picture of different reference graph coloring. Aiming at this problem, based on the instance of color image segmentation and image fusion technology, the use of deep learning is proposed to implement regional mixed color more and master the method. It can be divided into foreground color based on reference picture and background color based on prior knowledge. In order to identify multiple objects and background areas in the image and fuse the final coloring results together, a method of image coloring based on CNN is proposed in this paper. Firstly, CNN is used to extract their semantic information, respectively. According to the extractive semantic information, the color of the designated area of the reference image is transferred to the designated area of the grayscale image. During the transformation, images combined with semantic information are input into CNN model to obtain the content feature map of grayscale image and the style feature map of reference image. Then, a random noise map is iterated to make the noise map approach the content feature map as a whole and the specific target region approach the designated area of the style feature map. Experimental results show that the proposed method has good effect on image coloring and has great advantages in network volume and coloring effect.

1. Introduction

With the emergence of digital media technology and the popularity of the Internet, the animation industry has been greatly developed and advanced [1–3]. Animation works usually have two forms of expression, two-dimensional animation and three-dimensional animation, among which two-dimensional animation works have strong representation, character drawing and coloring are more natural, and three-dimensional animation works are not limited by the physical engine. At present, two-dimensional animation works still have a wide influence. Generally, ordinary animation video requires at least 25 frames per second to ensure the continuity of the video, while a 25-minute animation video requires 37,500 frames of images [4, 5]. Although the drawing of intermediate frames can be made according to the

reference of key frames, the heavy task still requires the cooperation of multiple workers. In addition, after the middle frame line draft image is completed by the ordinary painter, it should be checked and modified by the animation instructor to maintain the consistency of the middle frame action and color and ensure the continuous effect of the character action in the line draft video [6]. Therefore, the research on the coloring and auxiliary rendering of animation line draft image can not only help the new artist to improve the drawing efficiency but also to reduce the manpower and material resources required for drawing line draft and coloring. In general, the key frame usually refers to the first and last frame of the animation shot. The post-production mainly completes the synthesis of characters and backgrounds, the addition of light and shadow effects, and the work of film editing and dubbing [7, 8].

Coloring is a very important stage after animation line draft image creation, time-consuming, and tedious; current cartoon makers will adopt some business equipment and software speed line art image color work but did not greatly improve production efficiency. In these papers, a new automatic coloring method for line art images is proposed based on the reference of color images, and the method is extended to similar color areas [9–12]. Inspired by the successful application of generative models in image synthesis tasks in recent years, the researchers use deep convolutional neural networks (DCNNs) and put forward many methods for automatic coloring of line draft images, but the coloring results of these methods are not controllable and often accompanied by color artifacts. In recent years, thanks to the prosperity of Internet technology, digital media industry has also become the core industry of the 21st century knowledge economy, such as film and television advertising online games, and a film and television work or an online game to attract people's attention often needs a dazzling poster [13–15]. Good works are not only reflected in content design but also in color collocation. Whether for the creation of pictures or videos, colorization is an extremely important link. However, this is not an easy task; the choice and collocation of color are a test of the artist's artistic foundation and time-consuming. Moreover, for a good work that has been created, if the color of one of the objects is not satisfied with the need to recolor, the existing method is direct gray recolor, which is a huge project. Therefore, the multiarea coloring of images is a significant research work for both academia and industry [16, 17].

Convolutional neural network is a simple and efficient network completely different from previous deep learning models. General deep learning models contain an overall neural network, but CNN breaks this structural mode. It consists of two subnetworks, namely, generator subnetwork and discriminator subnetwork. The generator can be used to extract image features and generate false images. Discriminators can be used to discriminate between real and fake images, giving a probability to conclude that the image is more likely to be real or to generate a fake image. In this process, both the generator model and the discriminator model are constantly trained [18]. With the increase in iteration times, the generator model's false image forgery ability becomes stronger and stronger, and the discriminator's accurate identification ability of true and false images becomes stronger and stronger and finally tends to converge. Therefore, the CNN model is widely used in image processing and is one of the most commonly used models in the field of image coloring [19–22].

2. Related Works

An et al. [23] asked users to draw a color curve for graffiti and set the gradient range of the curve to control the spread of graffiti. Specifically, the method takes a set of diffusion curves as constraints and obtains the final image by solving Poisson's equation. However, all the above methods require a lot of manual interaction to achieve target coloring. In order to reduce the manual work and realize the specified

color style coloring, researchers proposed a coloring method based on reference image. JWA et al. [24] used graph structure to represent the relationship between different regions of line draft image and solved the matching problem through quadratic programming. However, complex line draft image is usually difficult to be accurately segmented, and the same semantic region will be divided into multiple blocks. At present, researchers have proposed line draft map guided by reference image based on the deep learning image coloring method which avoids the requirement of image segmentation accuracy. Chen et al. [25] used conditional generative adversarial networks (cGANs) to color grayscale images without requiring users to interactively fine-tune the coloring results. However, this method is only suitable for learning the relationship between grayscale and color image, not line image. Active learning framework learns domain classification labels in small data sets and helps users to select the data to be labeled in unlabeled sets, so as to continuously update the learning model parameters and improve the accuracy of the model for unlabeled region classification labels. Zeng et al. [26] proposed an adaptive active learning method, which combined information density calculation with least uncertainty calculation to select marked instances, different from previous methods of selecting marked data based on uncertainty. Farid et al. [27] proposed the coloring task which involves specifying a three-dimensional information, such as RGB channel, from the one-dimensional information of grayscale image, that is, intensity or brightness. The mapping between the one-dimensional information and three-dimensional information is not unique. Colorization is ambiguous in nature, and appropriate external information needs to be provided. Therefore, the coloring algorithm based on image brightness weighted color mixing and fast feature space distance calculation can achieve high quality static image at the cost of a small part of the calculation cost and improve the speed of the algorithm.

Kotecha et al. [28] trained an automatic system for colorization of black and white images and trained the model to predict the color information of every pixel in the black and white images by using deep network to learn the detailed features of the color images. Berger et al. [29] proposed a new automatic coloring method for comics. Image features include global features and local features. Global features include the overall outline of the image, while local features include some details of the image. Based on convolutional neural network, the network contains two subnetworks, local feature extraction network and global feature extraction network, in order to achieve the purpose of cartoon image processing at any resolution. In recent years, with the popularity of deep learning, the mainstream of cartoon coloring gradually developed is to use two different methods to realize simple coloring algorithm and deep learning model. Among the methods based on deep learning, a variety of deep learning models have been used to complete the task of image translation. Many researchers have tried to use human intelligence to solve the task of automatic coloring of pictures. A Thakur et al. [30] pointed out that there are few papers on image processing using unsupervised learning CNNs network, so they proposed

DCGAN, a deep volume network, to realize CNN's attempts in supervised learning and unsupervised learning, respectively. How to accurately cut the image region is also a major factor affecting the final image color accuracy and quality; for image region segmentation, many researchers have done related research. Oladi et al. [31] decided to consider not only adjacent pixels with similar intensity but also distant pixels with the same texture, so they combined the two to enhance the visual effect. Through experiments, they found that better results can be obtained when coloring pixels near edges based on texture similarity and pixels in smooth regions based on intensity similarity. The method can also be used to color comics, and they have developed a set of interface tools that allow users to tag and color and modify target images. Qiao et al. [32] implemented two coloring methods based on U-NET. The main innovations of this method are as follows: one is to train a deep neural network to directly predict the mapping from grayscale images with colored points to color images; second, the network will also provide users with a data-driven color palette, suggesting the ideal color of the gray map in a given location. This approach can also bring the benefit of reducing the workload for users, and it can also calculate the global histogram of a color reference map to color the gray map [10, 33].

From the above analysis, we know that the above methods have studied the automatic gray image coloring widely. However, some problem still exists. For example, no scholar has applied the CNN model to this field till now, so the research here is still a blank, which has great theoretical research and practical application value for logistics enterprises [34].

This paper consists of five parts. The first and second parts give the research status and background. The third part is the automatic gray image coloring by the CNN model. The fourth part shows the experimental results. The experimental results of this paper are introduced and compared and analyzed with relevant comparison algorithms. Finally, the fifth part gives the conclusion of this paper.

3. Automatic Gray Image Coloring by CNN Model

3.1. The Process of Automatic Gray Image Coloring. For any enterprise, capital is the source of its life; financing is a way to revitalize the enterprise capital, improve the effective utilization rate of capital, and obtain profits. With the modern new production organization mode—the new financing mode produced by supply chain—supply chain finance financing has become a hot spot, supply chain finance is called the general trend, and enterprises must have its reasons and conditions for supply chain finance financing. As a result, those with the greatest impact of stress and the greatest capacity to take the most drastic and effective action for change are the most likely to achieve the best performance. This paper will establish the supply chain financial performance evaluation index system of warehousing and logistics enterprises from the four dimensions of pressure dimension, action dimension, ability dimension, and driving factor dimension. The whole system of the method is given in Figure 1.

Except for the CNN model, VGG network achieved good results in ILSVRC positioning and classification tasks, respectively. VGG network inherited the main convolution-pooling network in AlexNet. They abandoned large convolution kernels and replaced them with multiple convolution kernels with a size of 3×3 , which could reduce the number of network parameters and increase the network depth. This can be said to be the in-depth version of AlexNet; deeper network better fits complex nonlinear problems. Even so, the number of VGG network parameters is still very large. Generally speaking, a VGG network contains 500 parameters, so the model takes up a lot of storage space. However, thanks to its excellent feature extraction ability, it is very suitable for the auxiliary task of feature extraction in some image processing tasks. Differential network does not refer to a specific network, but a structure that can be used in any network model. Residual structure is a connection mode that prevents network degradation through a hop connection. In addition, even with the use of ReLU activation function, the phenomenon of gradient disappearance will occur with the increase in the number of network layers, while the residual structure can solve the above problems. The residual structure adopts the method of skip connection in the network structure. In conclusion, the CNN model shows better performance than the VGG network; hence, the CNN is selected in this paper.

3.2. Convolutional Neural Network. In the process of image processing, we often use matrix convolution to calculate the feature of image. There are two types of matrix convolution: full convolution and valid convolution. The definition of full convolution is as follows:

$$\begin{aligned} z(u, v) &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x_{i,j} \cdot k_{u-i,v-j}, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2. \end{aligned} \quad (1)$$

Assuming that X is the m -order matrix and k is the n -order matrix, the definition of effective convolution is

$$z(u, v) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x_{i+u,j+v} \cdot k_{roti,j} \cdot \chi(i, j). \quad (2)$$

Assume that

$$\chi(i, j) = 0 \text{ or } 1. \quad (3)$$

Convolution layer (the previous layer is the input layer): in the convolution, the data is input to the input layer in 3d form, and then the convolution kernel of the first layer and the corresponding functional module convolve the input data. We add a bias term to each output, and the output of the convolution layer is as follows:

$$\begin{aligned} z_{u,v}^{(l)} &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x_{i+u,j+v}^{(l-1)} \cdot k_{rot}^{(l),j} \cdot \chi(i, j) + b^{(l)}, \\ \chi(i, j) &= 0 \text{ or } 1, \\ a_{u,v}^{(l)} &= f(z_{u,v}^{(l)}). \end{aligned} \quad (4)$$

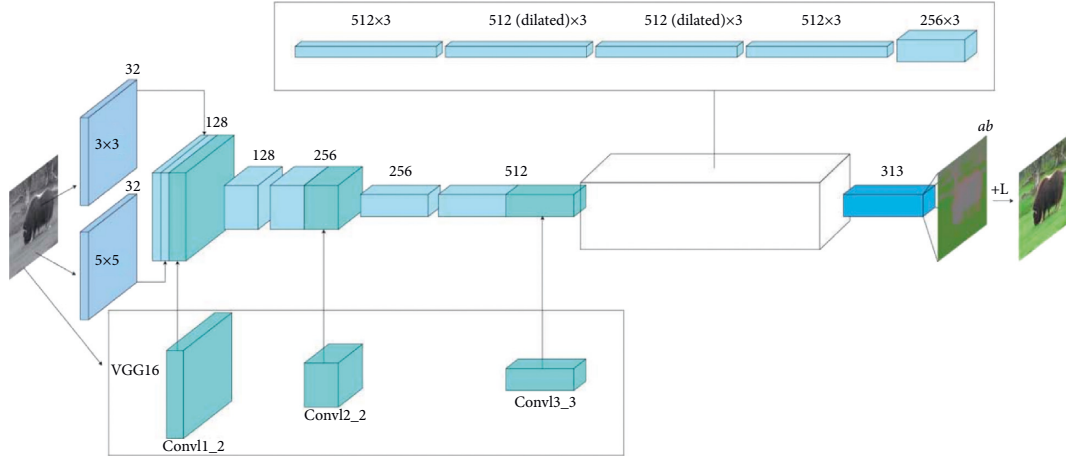


FIGURE 1: The framework and flow chart of the proposed method.

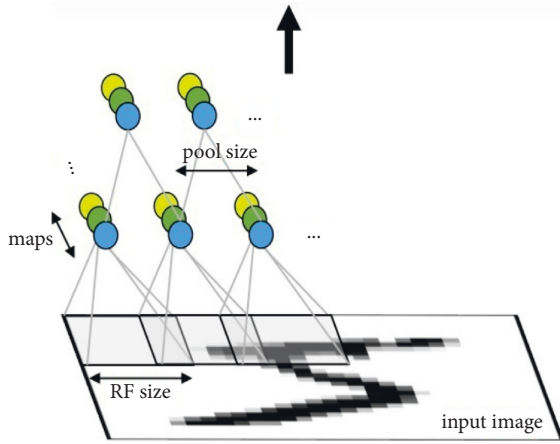


FIGURE 2: The schematic diagram of convolutional neural network.

To ensure that variance is positive, that is, variance exists and is finite as follows:

$$\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1. \quad (5)$$

After the input feature graph passes through the convolution layer, we get its feature graph. Now, we hope to use these feature graphs to train the classifier. Theoretically, we can use all the extracted feature graphs to train the classifier. In order to solve this problem, we can use the aggregation of the statistical method. For example, we can replace all the original features with the average of the image features, which is faster and less prone to over-fitting than all the features used. So, this aggregation is called pooling, and pooling is divided into average pooling and max pooling. Here, we take the average pooling method as an example and use the weight of each unit of the convolution kernel. After each convolution operation, a bias unit is still added. The output of the subsampling layer is as follows:

Image No.	Origin Image	Gray Image	Reference Image	Colorized Image
1				
2				
3				
4				
5				
6				
7				

FIGURE 3: Coloring results for different categories of images.

$$z_{i,j}^{(l+1)} = \beta^{(l+1)} \sum_{u=ir}^{(i+1)r-1} \sum_{v=jr}^{(j+1)r-1} a_{u,v}^{(l)} + b^{(l+1)},$$

$$a_{i,j}^{(l+1)} = f(z_{i,j}^{(l+1)}), \quad (6)$$

$$K_a^\phi = \frac{3 - 3(\alpha_1 + \beta_1)^2}{1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2} - 3.$$

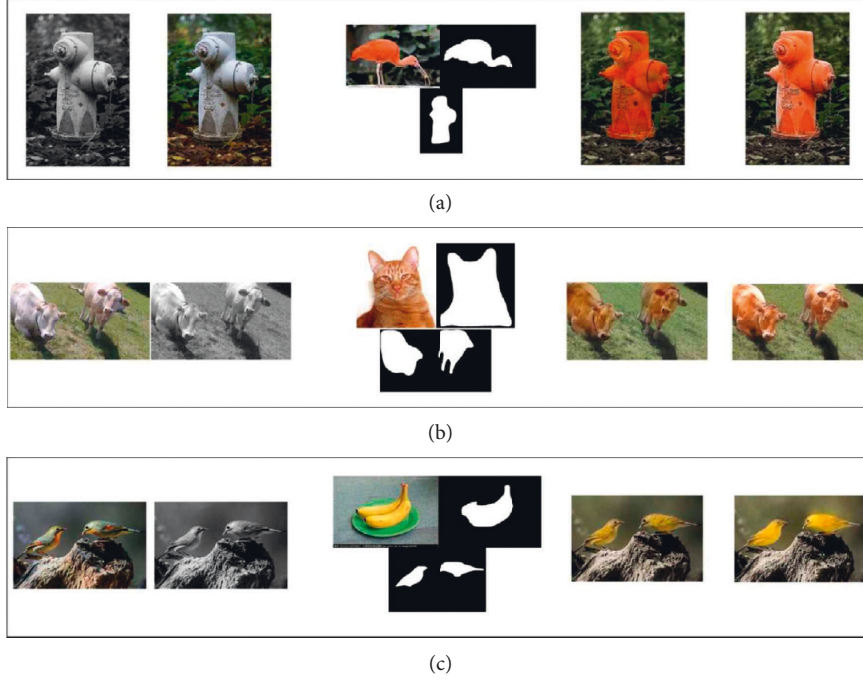


FIGURE 4: Category at the same time color result comparison.

If the subsampling layer is followed by the convolution layer, the calculation method is the same as that described by the multilayer neural network, and the output is as follows:

$$z_{u,v}^{(l+2)} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} a_{i+u,j+v}^{(l+1)} \cdot k_{i,j}^{(l+2)} \cdot \chi(i,j) + b^{(l+2)}, \quad (7)$$

$$a_{u,v}^{(l+2)} = f(z_{u,v}^{(l+2)}).$$

The schematic diagram of a typical convolutional neural network is shown in Figure 2:

4. Experimental Results and Analysis

4.1. Introduction to Experimental Environment and Data Set. This line of business on Windows 10 OS runs a HMP RTX 2070S 8 GB video memory, AMD Ryzen 2400G CPU, 16 GB DDR4 software environment: Deep learning framework Pytorch 1.8. Python 3.7 CUDA 10.0 data set uses Place 365 outdoor landscape, including buildings, cabins, landscapes, courtyards, and more than 50 categories. The Epoch is 10. The comparison algorithm is trained by a business line with the same data set, and the Epoch trained is also 10.

In this experiment, the initial learning rate is set to 0.02 by experience, the attenuation weight coefficient is 0.0001, the updated weight is 0.1, the updated weight attenuation system is 0.0002, the maximum number of iterations is 10000, epoch is 600 times, and the random gradient descent method Batch is selected. Batch training options are 50.

4.2. Experimental Results Analysis. Firstly, the comparison of the results of the real standard foreground coloring using random noise graph and image conversion network,

respectively, includes the comparison of its coloring efficiency, and then the two foreground coloring methods, respectively, combined with U-NET and Poisson fusion, are shown to achieve the whole picture coloring. And the multiregion color rendering method based on random generated noise graph and basis forward transformation network is proposed, respectively. Figure 3 shows the coloring effects of different categories of images.

Transformation of regional colours is possible in random noise images and image transformation networks, but there are massive differences between the two speeds and the colours of the ginseng, and each colour needs to be trained, while the image transformation network taken is made up of cyclic transformations. The network generates the noise graph, then updates the parameters of the image conversion network training, and saves the optimal solution.

When one or more shells appear in the picture, you can select a different reference image for each earmark in the picture and color all ICONS in the picture at the same time by selecting the reference image. According to the input semantic map of gray image, one or more marks in the image are colored. In addition, due to the addition of semantic information as a strong constraint condition, the two color labeling methods in this paper have stronger constraints, so as to obtain better graphic coloring effect. Figure 4 shows the coloring results of the same and different categories, which can make the coloring results more diverse and conducive to the user's image creation. The results show the combination of foreground coloring, back coloring, and Poisson melting.

It can be seen that the simple fusion effect directly depends on the result of image instance segmentation, and the quality of the result of instance segmentation determines the fusion effect. However, the current instance



FIGURE 5: Different methods for recoloring color pictures.



FIGURE 6: Different methods for recoloring gray image.

segmentation technology can only circle the general target, and there is still a lack of edge processing. Therefore, this paper uses the CNN algorithm to fuse the background and foreground after coloring, so that the edge can smoothly transition.

Because the final effect of coloring is difficult to measure with mathematical way, the richer the color, the greater the final loss, but the final effect is acceptable. Figures 5 and 6 show the comparison of some experimental effects. Figure 6 shows the recoloring effect of color photos, and Figure 6 shows the coloring effect of black and white photos. From the perspective of recoloring effect of color images, the images colored by the proposed method are more colorful and have better processing of details and light and shadow effects than other algorithms. DCGAN's effect is always

dark, and the algorithm in this paper does not recognize the ground in group E , but other groups of images have relatively good effects.

In the colorful image coloring shown in Figure 5, the method presented in this paper has good semantic properties, vivid colors, and good restoration of sky. The BP algorithm is also good in general, but the restoration of sky is not accurate, and RNN is dull. The algorithm in this paper has a better coloring effect at the gap between leaves and sky and is more accurate in coloring buildings, while the BP algorithm is green. The proposed algorithm is relatively accurate to color the ground, while the BP algorithm may treat the map as the ocean. The coloring effect of several groups of results of the RNN algorithm is too dull.

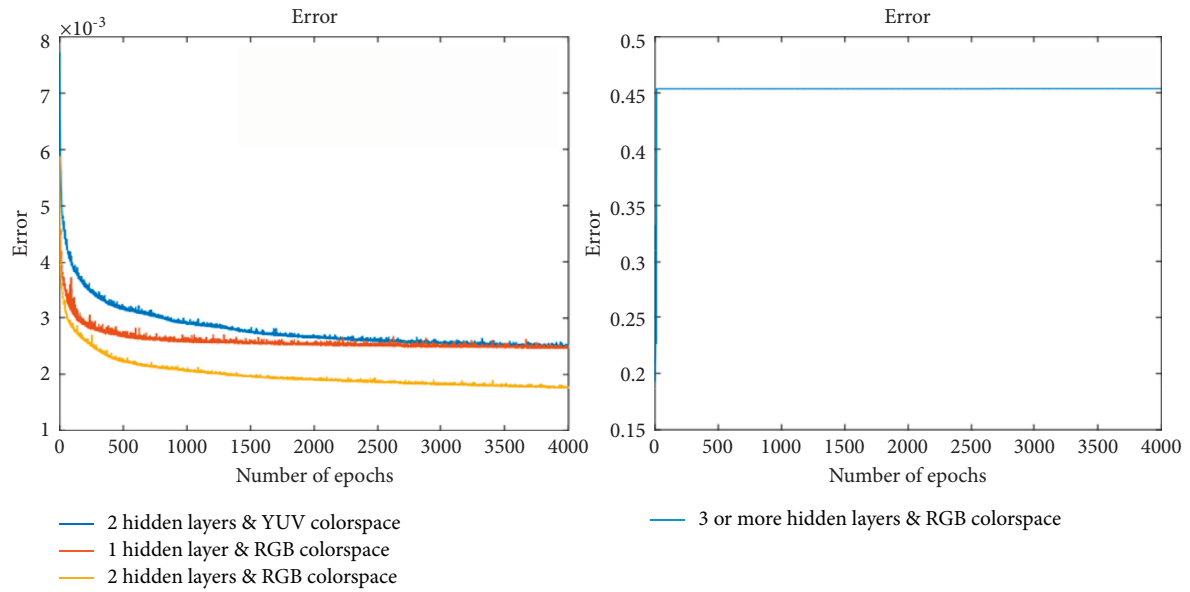


FIGURE 7: The error curves of different scenes by CNN.

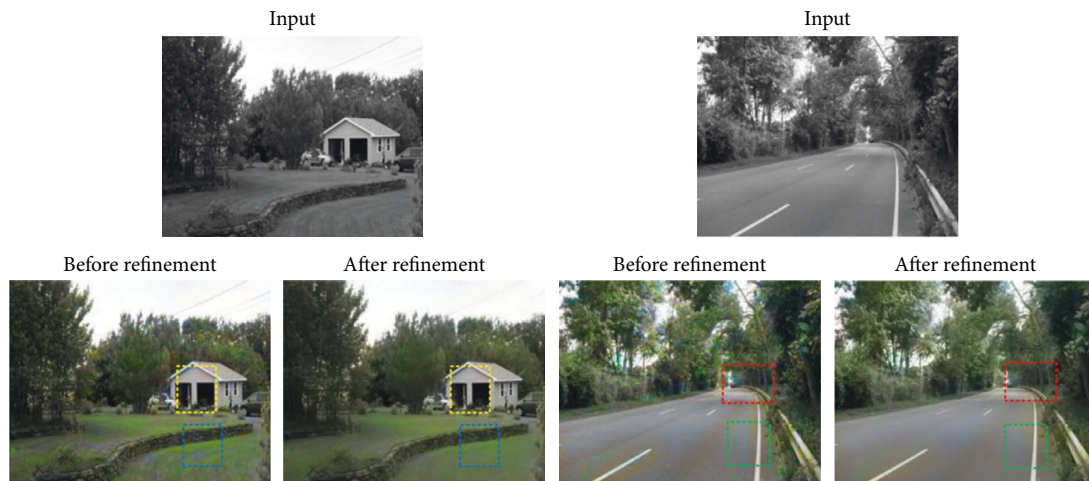


FIGURE 8: Effect comparison before and after color refinement.

As shown in Figure 6, it can be seen from the comparison with the original picture that the color of CNN model proposed in this paper is more realistic and natural; that is to say, this paper has obtained the best image coloring effect. .

As shown in Figure 7, just add some color prompt graffiti lines to each part of the scene, and the model can render other uncolored parts in the image area according to the prompts. Moreover, this color rendering is not pure color filling. The transition of light and dark colors makes the whole coloring effect more natural and not rigid. For example, in the pictures of the second group of buildings, only a pure blue graffiti hint line is given instead of a pure color. However, the interaction between colors is not very obvious there are some rise in space, such as in the first set of interaction diagrams had two yellow graffiti decorate tip lines, in addition to other part is not in the yellow line, but the generated version of athletes central sleeve are also be

rendered into yellow, with the expected painting effect exists some gap. However, MSE generally showed a downward trend.

Figure 8 shows the effect comparison before and after color refinement. In the CNN algorithm, each gray pixel performs full-image search and matching, and the pixel with the smallest error is selected in the source image for matching. Therefore, all pixels in the target image should be able to find matching pixels in the source image. However, we are still unable to obtain a very satisfactory method for both color quality and color speed. How to make before and after color refinement achieve a win-win situation is the ultimate goal of this study. In terms of improving quality, more and more matching guiding factors have been proposed, instead of relying only on the brightness mean and brightness variance as the only criteria for matching search. For example, the slope and kurtosis containing nonlinear

properties of images can guide pixel matching. In order to improve the speed, the tree structure is used to classify pixels as much as possible, so that the matching process has a clear goal, rather than blindly carrying out full image search.

5. Conclusions

In recent years, with the vigorous development of computer hardware equipment, artificial intelligence has appeared more and more frequently in people's vision. The image automatic color not only can increase the interest in everyday life but also can improve the efficiency of some work, such as poster making, etc. So, the gray image coloring technology has also attracted the attention of many researchers. In this article through the large number of literatures, the survey found existing in coloring technology can meet the needs of most of the color but still exist with some disadvantages, such as not to picture one or more specific areas in color and not only change the image of one or more specific areas in color. Aiming at these deficiencies, this paper uses semantic and image fusion technology to realize the image segmentation of regional color more, separately from the multiple target and the background color of the image. Background coloring is carried out end-to-end coloring by CNN, while target coloring needs to be colored according to the color reference graph. Target coloring is divided into two coloring methods, one is iterative image coloring and the other is training image conversion.

Gray image automatically chromatically is an important research direction in the field of image. With the continuous development of deep learning in recent years, automatic coloring of gray image is gradually realized based on the deep learning model. The simple coloring algorithm is generally inferior to the deep learning model. Both interactive coloring reference coloring and automatic coloring can be realized by deep learning. However, there are still some limitations in using deep learning to complete the gray image automatic coloring task; for example, the color effect learning is not in place, the gray image line contour recognition is not accurate evaluation method and is not unified, and there is no inclusive professional gray image coloring platform and so on.

Using the color based on classification of network, this paper proposes a new CNN network using a traditional Gaussian convolution encoder and a hollow convolution stack structure to perform automatic coloring of gray images. Compared with the results of other business methods, it has advantages in the final coloring effect and volume control.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] N. K. E. Abbadi and E. S. Razaq, "Automatic gray images colorization based on lab color space," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 3, pp. 1501–1509, 2020.
- [2] M. Xu and Y. Ding, "Fully automatic image colorization based on semantic segmentation technology," *PLoS One*, vol. 16, no. 11, Article ID e0259953, 2021.
- [3] D. Wu, J. Gan, J. Zhou, J. Wang, and W. Gao, "Fine-grained semantic ethnic costume high-resolution image colorization with conditional GAN," *International Journal of Intelligent Systems*, vol. 37, no. 5, pp. 2952–2968, 2022.
- [4] N. K. El Abbadi and E. Saleem, "Image colorization based on GSVD and YCbCr color space," *Kuwait Journal of Science*, vol. 46, no. 4, pp. 47–57, 2019.
- [5] Y. Xiao, A. Jiang, C. Liu, and M. Wang, "Semantic-aware automatic image colorization via unpaired cycle-consistent self-supervised network," *International Journal of Intelligent Systems*, vol. 37, no. 2, pp. 1222–1238, 2022.
- [6] E. Saleem and N. K. El Abbadi, "Auto colorization of gray-scale image using YCbCr color space," *Iraqi Journal of Science*, vol. 61, pp. 3379–3386, 2020.
- [7] L. C. Muskat, Y. Kerkhoff, P. Humbert, T. W. Nattkemper, J. Eilenberg, and A. V. Patel, "Image analysis-based quantification of fungal sporulation by automatic conidia counting and gray value correlation," *MethodsX*, vol. 8, Article ID 101218, 2021.
- [8] P. Zheng, "Multisensor Feature Fusion-Based Model for Business English Translation," *Scientific Programming*, vol. 2022, Article ID 3102337, 10 pages, 2022.
- [9] A. Bouida, M. Beladgham, A. Bassou, and I. Benyahia, "Quality and texture analysis of biometric images compressed with second-generation wavelet transforms and SPIHT-Z encoder," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 3, pp. 1325–1339, 2020.
- [10] X. Bi, W. Yao, Z. Zhang, S. Huang, J. Liu, and B. Chen, "Image steganography algorithm based on image colorization," in *Proceedings of the International conference on signal image processing and communication (ICSIPC 2021)*, vol. 11848, Article ID 1184818, Chengdu, Sichuan, China, April 2021.
- [11] W.-T. You, L.-Y. Sun, Z.-Y. Yang, and C.-Y. Yang, "Automatic advertising image color design incorporating a visual color analyzer," *Journal of Computer Languages*, vol. 55, Article ID 100910, 2019.
- [12] B. Liu, J. Gan, B. Wen, Y. LiuFu, and W. Gao, "An automatic coloring method for ethnic costume sketches based on generative adversarial networks," *Applied Soft Computing*, vol. 98, Article ID 106786, 2021.
- [13] M. Zare, K. B. Lari, M. Jampour, and P. Shamsinejad, "Multi-GANs and its application for Pseudo-Coloring," in *Proceedings of the 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA)*, pp. 1–6, IEEE, Tehran, Iran, March 2019.
- [14] G. Kong, H. Tian, X. Duan, and H. Long, "Adversarial edge-aware image colorization with semantic segmentation," *IEEE Access*, vol. 9, pp. 28194–28203, 2021.
- [15] C. Liu, L. Cao, and K. Du, "Colorization of characters based on the generative adversarial network," *Cloud Computing, Smart Grid and Innovative Frontiers in Telecommunications*, pp. 636–649, Springer, Heidelberg, Germany, 2019.
- [16] S. Wan, Y. Xia, L. Qi, Y.-H. Yang, and M. Atiquzzaman, "Automated colorization of a grayscale image with seed points

- propagation,” *IEEE Transactions on Multimedia*, vol. 22, no. 7, pp. 1756–1768, 2020.
- [17] S. Huang, C. Lin, K. Zhou, Y. Yao, H. Lu, and F. Zhu, “Identifying physical-layer attacks for IoT security: an automatic modulation classification approach using multi-module fusion neural network,” *Physical Communication*, vol. 43, Article ID 101180, 2020.
- [18] J. Zhang, S. Zhu, K. Liu, and X. Liu, “UGSC-GAN: User-guided Sketch Colorization with Deep Convolution Generative Adversarial networks,” *Computer Animation and Virtual Worlds*, Article ID e2032, 2021.
- [19] N. Zhang, P. Qin, J. Zeng, and Y. Song, “Image colorization algorithm based on dense neural network,” *International Journal of Performability Engineering*, vol. 15, no. 1, pp. 270–280, 2019.
- [20] P. Jain and U. Ghanekar, “Robust watermarking technique for textured images,” *Procedia Computer Science*, vol. 125, pp. 179–186, 2018.
- [21] B. N. Mohammed and H. B. Ahmad, “Advanced car-parking security platform using Arduino along with automatic license and number recognition,” *Academic Journal of Nawroz University*, vol. 10, no. 1, pp. 1–6, 2021.
- [22] L. Chen, J. Han, and F. Tian, “Colorization of fusion image of infrared and visible images based on parallel generative adversarial network approach,” *Journal of Intelligent and Fuzzy Systems*, vol. 41, no. 1, pp. 2255–2264, 2021.
- [23] J. An, K. G. Kpeyton, and Q. Shi, “Grayscale images colorization with convolutional neural networks,” *Soft Computing*, vol. 24, no. 7, pp. 4751–4758, 2020.
- [24] M. Jwa and M. Kang, “Grayscale image colorization using a convolutional neural network [j],” *Journal of the Korean Society for Industrial and Applied Mathematics*, vol. 25, no. 2, pp. 26–38, 2021.
- [25] Y. Chen, Y. Luo, Y. Ding, and B. Yu, “Automatic Colorization of Images from Chinese Black and white Films Based on cnn,” in *Proceedings of the 2018 International Conference on Audio, Language and Image Processing (ICALIP)*, pp. 97–102, IEEE, Shanghai China, July 2018.
- [26] X. Zeng, S. Tong, Y. Lu, L. Xu, and Z. Huang, “Adaptive medical image deep color perception algorithm,” *IEEE Access*, vol. 8, pp. 56559–56571, 2020.
- [27] M. S. Farid, M. Lucenteforte, and M. Grangetto, “Evaluating virtual image quality using the side-views information fusion and depth maps,” *Information Fusion*, vol. 43, pp. 47–56, 2018.
- [28] K. Kotecha, “Automatic image colorization using GANs,” in *Proceedings of the Soft computing and its engineering applications: second international conference, IcSoftComp 2020*, vol. 1374, p. 15, Springer Nature, Changa, India, December, 2020.
- [29] D. R. Berger, H. S. Seung, and J. W. Lichtman, “VAST (volume Annotation and segmentation tool): efficient manual and semi-automatic labeling of large 3D image stacks,” *Frontiers in Neural Circuits*, vol. 12, p. 88, 2018.
- [30] G. K. Thakur, B. Priya, and R. K. Mishra, “An efficient coloring algorithm for time detracton of sign image segmentation based on fuzzy graph theory,” *Journal of Applied Security Research*, vol. 14, no. 2, pp. 210–226, 2019.
- [31] M. Oladi, A. Ghazilou, S. Rouzbehani, N. Zarei Polgardani, K. Kor, and H. Ershadifar, “Photographic application of the Coral Health Chart in turbid environments: the efficiency of image enhancement and restoration methods,” *Journal of Experimental Marine Biology and Ecology*, vol. 547, Article ID 151676, 2022.
- [32] T. Qiao, R. Shi, X. Luo, M. Xu, N. Zheng, and Y. Wu, “Statistical model-based detector via texture weight map: application in re-sampling authentication,” *IEEE Transactions on Multimedia*, vol. 21, no. 5, pp. 1077–1092, 2018.
- [33] D. Ouyang, R. Furuta, Y. Shimizu, Y. Taniguchi, R. Hinami, and S. Ishiwatari, “Interactive Manga Colorization with Fast Flat Coloring,” in *Proceedings of the SIGGRAPH Asia 2021 Posters*, pp. 1–2, Tokyo, Japan, December 2021.
- [34] S. Huang, X. Jin, Q. Jiang et al., “A Fully-Automatic Image Colorization Scheme Using Improved CycleGAN with Skip connections,” *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26465–26492, 2021.

Research Article

A Multitask Deep Learning Framework for DNER

Ran Jin ¹, Tengda Hou ¹, Tongrui Yu ¹, Min Luo ², and Haoliang Hu ¹

¹College of Big Data and Software Engineering, Zhejiang Wanli University, No. 8 South Qianhu Road, Ningbo, China

²Ningbo University of Finance & Economics, No. 899 College Road, Ningbo, China

Correspondence should be addressed to Min Luo; luomin@nbufe.edu.cn and Haoliang Hu; huhaoliang79@163.com

Received 5 February 2022; Accepted 18 March 2022; Published 16 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Ran Jin 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.

Over the years, the explosive growth of drug-related text information has resulted in heavy loads of work for manual data processing. However, the domain knowledge hidden is believed to be crucial to biomedical research and applications. In this article, the multi-DTR model that can accurately recognize drug-specific name by joint modeling of DNER and DNEN was proposed. Character features were extracted by CNN out of the input text, and the context-sensitive word vectors were obtained using ELMo. Next, the pretrained biomedical words were embedded into BiLSTM-CRF and the output labels were interacted to update the task parameters until DNER and DNEN would support each other. The proposed method was found with better performance on the DDI2011 and DDI2013 datasets.

1. Introduction

With the rapid development of biomedicine and the exponential growth of publications have made it hard to extract a number of drug-related information. It is essential to extract valuable information if we want to make the best of medical text. Medicine is a class of chemical substances that are highly associated with biological research. It is of vital significance to observe how to accurately capture the entity information as contained in medicine. Drug refers to chemical name, generic term, or brand name. As a chemical product usually has a complex name, the brand name may not exactly identify a drug in the expiry of relevant patents. For example, the drug “quetiapine” is associated with the brand name “Seroquel XR.” Therefore, a special generic term, which needs to be explicitly defined for drug approval, should be designed for standard scientific reports and labels. Drug-specific names are subject to tight control by WHO (World Health Organization) and some organizations in the USA and elsewhere. For example, the European Medicines Agency (EMA) finalized the naming scheme fit to drug function for ease of pronunciation and translation and developed some criteria that differentiate a drug name from others so as to avoid any transcription and

replication error in the R&D process [1]. This would justify the automatic extraction of potential medical information from massive biomedicine-related publications as a crucial part of biomedical research and industrial medicine manufacturing.

Drug-Named Entity Recognition (DNER), which is intended to identify the drugs referred to in unstructured drug texts, is an underlying task of recognizing the span and type of the named entity subordinated to predefined semantic types. Unlike ordinary NERs (Named Entity Recognition), DNER generally consists of long label sequences and contains plenty of alternate spellings of synonyms and entities, resulting in the inefficiency of drug dictionary and hard detection of entity boundaries. In this regard, Drug-Named Entity Normalization (DNEN) is also believed to be a crucial task.

DNEN, which is intended to map the acquired DNERs to a controlled vocabulary, is usually considered a task subsequent to DNER. Both DNEN and DNER can be deemed as sequence labeling problems. Figure 1 illustrates an example with respect to DNER and DNEN tasks, the input text contains the drug-specific name “Omeprazole” and the R&D organization “Astra Pharmaceuticals”, and the label of each word in the text and its entity ID are output.

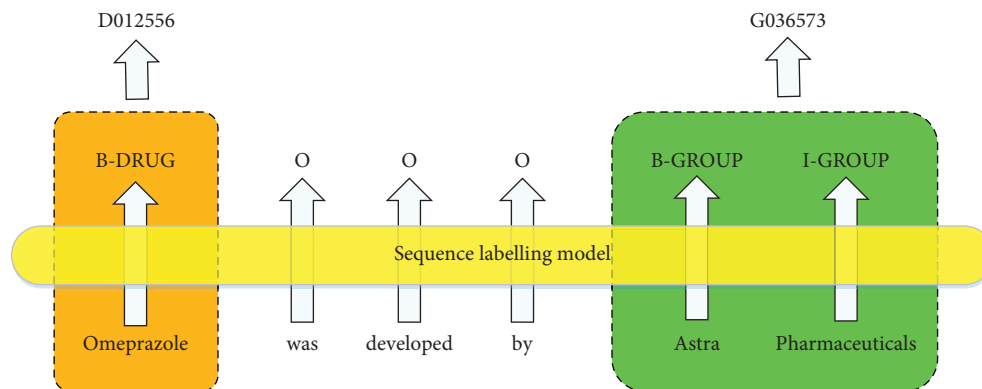


FIGURE 1: An example of DNER and DNEN tasks.

As the naming scheme, evaluation criteria and cross-border synchronization have been developing dynamically for many years, and there is no definitive dictionary or grammar applicable to drug names; DNER and DNEN processes are subject to many challenges: (1) the rapid updates of drug-related knowledge make it hard for a handmade dictionary to meet actual needs; (2) language tends to be complex and there is a scarcity of high-quality label texts; (3) the simple modeling of DNER and DNEN cannot allow both processes to support each other.

It is intended that the proposed model can capture more resourceful semantic features and identify the representation of polysemous and ambiguous words in drug sequence, thus accurately recognizing drug names. A multitask deep learning model multi-DTR (Multi-Drug Tip Recognition) was proposed, and the principal contributions of this work were that text information can be exploited by extracting the character-level representations of words, embedding words based on biomedicine pretraining, and extracting the features by context-sensitive word embedding after ELMo (Embeddings from Language Models) training. To make the best of the training data, a multitask learning strategy was taken, which allows for the explicit feedback of DNER and DNEN and makes different tasks support each other.

This article is structured as follows: in Section 2, some related works on DNER and DNEN were presented; in Section 3, the proposed neural network framework was described; in Section 4, relevant datasets and parameter setups were briefed; in Section 5, the result of the assessment was reported in particulars; in Section 3, a conclusion was drawn.

2. Related Works

NER is one of the underlying tasks in NLP, but there are a limited number of related works on DNER [2–4]. The access to some large-scale biomedical corpora [5–7] has enabled some generic NER models to be widely used in DNER. Common methods applicable to DNER can be roughly categorized into rule-based methods [8], dictionary-based methods [9], and machine learning-based methods [10]. In the case of rule-based methods, a number of labor resources are required to lay down rules, but the ambiguity and

variability of terms are overlooked. If the target text appears to be complex, rule-based methods are found with a low recognition rate [11]. Tsuruoka et al. [12] made use of logistic regression to learn string similarity measures from the dictionary and performed soft character matching to avoid large difference of association due to exact string matching. Hettne et al. [13] developed a rule-based method for term filtering and disambiguation, then merged dictionaries to recognize small molecules and drugs as contained in the text. Eriksson et al. [14] created a Danish dictionary to recognize Adverse Drug Event (ADE) that may potentially occur in unstructured clinical narrative text. Despite this, the actual application needs can hardly be met due to a lack of dictionary and rapid update of biomedicine terms. The machine learning-based NER is currently a prevailing research interest. Cocos et al. [15] used ZRNN coupled with pretrained word embedding to recognize ADE on Twitter. Zeng et al. [16] performed automatic searching of words and character-level features in drug texts on LSTM-CRF (Long Short-Term Memory-Conditional Random Field) structure. To date, BERT (Bidirectional Encoder Representations from Transformers) [17] is the great hit model in the sector of Natural Language Processing (NLP). In the case of BERT, a transformer encoder was used and the upper and lower layers of the model are fully connected by a self-attention mechanism so that text information can be better processed. Lee et al. [18] ran a large-scale pretraining in respect of BERT (treated as a basic model) on PubMed and PMC and then developed the BioBERT (Biomedical Bidirectional Encoder Representations from Transformers) model. Despite the extraordinary properties, this model caused an enormous consumption of hardware resources in the training process.

DNEN, also a key part of information extraction, is generally listed as a subtask [19, 20] for some biomedicine-related NLP assessment tasks. Kang et al. [21] normalized disease-specific names by constructing a symptom text model and performing a comparative analysis. Lee et al. [22] used a dictionary to look up and standardize the entity. Lou et al. [23] proposed a transition-based model applicable to the recognition and normalization of joint disease entity, but such model heavily relies on handmade features and task types.

3. Neural Network Framework

In this article, the character feature representations (e.g., amidopyrine, aminophenazone, and aminopyrine) of an input word were extracted through Convolutional Neural Networks (CNN). Next, the extracted character features and words were embedded and input to BiLSTM (Bidirectional Long Short-Term Memory). The two-way LSTM (Long Short-Term Memory) was used to capture two separate hidden states (forward and backward) of each sequence, obtain the context-sensitive information, then connect two hidden states until the final output is generated. In the final step, the output vector of BiLSTM was backfed to CRF for jointly modeling the label sequence. DNER and DNEN can give back to each other by the output of two tasks, reduce the load of calculations, and realize the enhancement effect of both tasks.

3.1. Embedded Layer. For deep mining of drug-related information in the input text, the features were extracted by pretrained word embedding, context-sensitive word embedding, and character embedding.

3.1.1. Pretrained Word Embedding. The rapid development of deep learning technology has led to an extensive use of word embedding, which offers an alternative to numerical representation of text (such as Word2Vec [24] and Glove [25]). Yu et al. [26] found that embedding pretrained words into unlabeled data would have many NLP tasks significantly improved. As inspired by Glove [25], we used the word representation method based on global word frequency statistics to pretrain data on PMC (PubMed Central) and PubMed biomedical corpora and to embed pretrained word vectors into the model.

3.1.2. Character Representation. Evidence has shown that character information is crucial for sequence labeling tasks [16, 27]. Colobert et al. [28] suggested that the integrity of words can be used to label words, and local features extracted by CNN are exploited to construct all feature vectors. Ling et al. [29] tried to use character-level two-way LSTM for POS labeling, but the result of the experiment indicates that the performance of character-level two-way LSTM highly resembles CNN, but a heavier load of calculations is requested. Santos et al. [30] was the first researcher who suggested using CNN to learn character-level representations of words and associate them with the representations of common words. A number of subsequent works [31, 32] supported that the word-level information (such as prefixes and suffixes) can be leveraged to the extent possible by character-based word representation. Zhao et al. [33] exploited attention-based CNN to capture the association between context-sensitive information and discontinuous words. Strubell et al. [34] proposed ID-CNN (Iterated Dilated Convolutional Neural Network) as the generally dilated CNN architecture that improves the computational efficiency to the extent possible. Chiu et al. [35] used CNN to extract character vectors of a

specific length from the word-specific characters, cascade them with the encoded features, then transmit them through the convolutional layer and the max layer.

In this article, CNN was used to acquire the character-level representation of a word. As is seen from Figure 2, the feature encoding process as a part in Chiu et al. [35] was deleted, the Dropout layer was added to prevent overfitting of CNN, and we finally had a word-specific character vector.

3.1.3. Context-Sensitive Word Embedding. ELMo, a language model based on features, can model words given the context. Unlike Word2Vec and other word sectors that use a simple lookup table to obtain the unique representation, the word sector in ELMo represents the function of the internal network state. Even for the same word, the word sector shows changes dynamically. Thus, it first adopts two-way LSTM for pretraining and the two-way concept of ELMo is reflected through the network structure, which comprises the forward LSTM model and the backward LSTM model. The construction of the model is shown in Figure 3.

ELMo comes with a task attribute and is a linear combination represented by the middle layer of biLM. With respect to a given word, biLM of a L layer can obtain the representation of $2L + 1$:

$$ELMo_k = \sum_{j=0}^L w h_{k,j}^{LM},$$

$$R_k = \left\{ x_k^{LM}, \overset{\rightarrow LM}{h}_{k,j}, \overset{\leftarrow LM}{h}_{k,j} \mid j = 1, \dots, L \right\} = \{h_{k,j}^{LM} \mid j = 0, \dots, L\},$$
(1)

where w is the weight of softmax-normalized, x_k^{LM} denotes the input initial word vector, $\overset{\rightarrow LM}{h}_{k,j}$ denotes the forward LSTM output, and $\overset{\leftarrow LM}{h}_{k,j}$ denotes the backward LSTM output. The context-sensitive dynamic word embedding as obtained from the above can more accurately reflect the complex semantic and grammatical features of the text.

3.2. Sequence Labeling. Some deficiencies of the character-level model include the multiple growth of the effective sequence size and a lack of inherent meaning in the characters. Thus, RNN can be used to process time series data of any length using neurons with self-feedback. However, it was reported [36] that RNN is usually inclined to the nearest input of the sequence in practice and cannot process long-term dependencies. Certain variants based on recurrent neural networks, such as Gated Recurrent Unit (GRU) and LSTM, have proven extraordinary performance. Yang et al. [37] used GRUs at the character- and word-level to encode morphological and context-sensitive information. Huang et al. [38] were the first researchers who used BiLSTM for sequence sorting and results showed that this model is less dependent on word embedding and can capture two hidden states (forward and backward) of each sequence well with strong robustness.

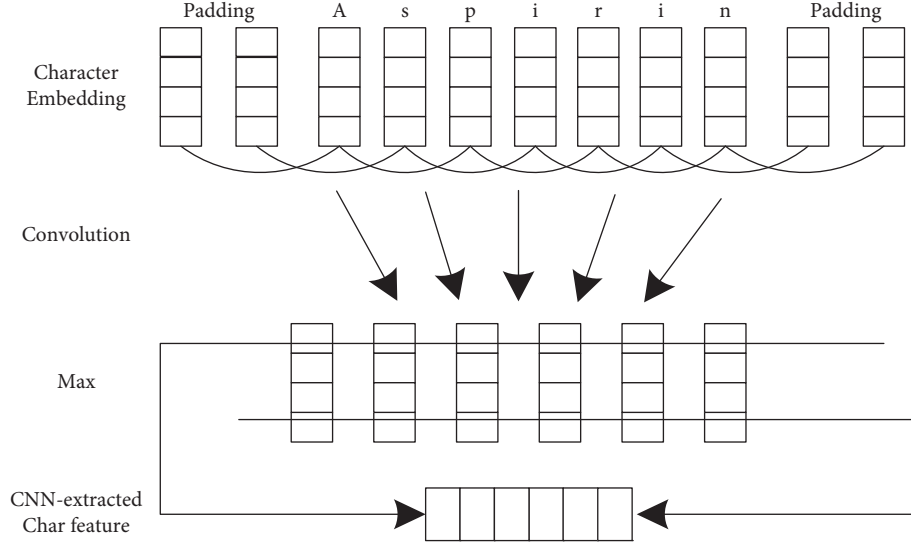


FIGURE 2: CNN used to extract a character-level representation of words.

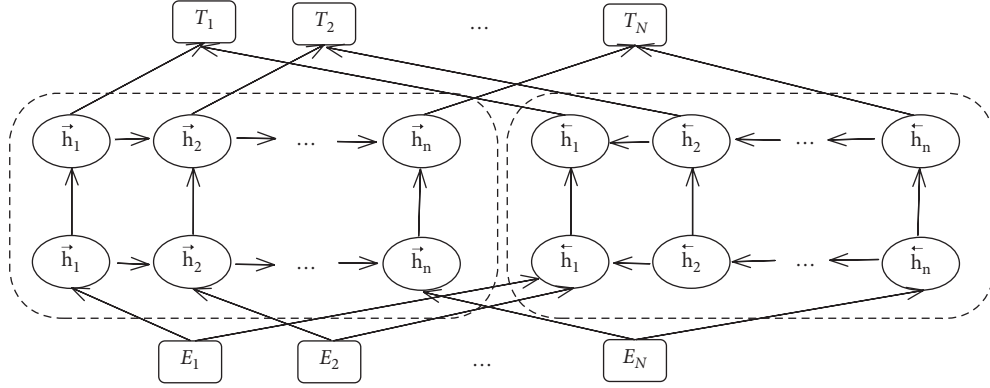


FIGURE 3: Structure of the ELMo model.

Both DNER and DNEN can be seen as sequence labeling tasks. In this work, BiLSTM was used to model the input character-level information, pretrained word embedding, and contextualized word embedding. It inputs a vector sequence containing n words (x_1, x_2, \dots, x_n) , then calculates the hidden state sequence (h_1, h_2, \dots, h_n) , and outputs the label (o_1, o_2, \dots, o_n) . Finally, the equation with respect to an update of the LSTM unit would be as follows:

$$\begin{aligned}
 i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i), \\
 c_t &= (1 - i_t) * c_{t-1} + i_t * \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c), \\
 o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o), \\
 h_t &= o_t * \tanh(c_t),
 \end{aligned} \tag{2}$$

where σ is elementwise sigmoid function, $*$ is elementwise product, x_t denotes the input vector at t , h_t is the hidden vector (also referred to as "output vector"), it denotes the value of the memory gate, c_t denotes the cell state, o_t denotes the value of the output gate, W_{xi} , W_{xc} , and W_{xo} denote the weight matrix of different gates of the input x_i , W_{hi} , W_{hc} , and W_{ho} are the weight matrix of the hidden state h_i , and b_i , b_c ,

and b_o denote the offset vector. Then the final output vector $h_t = [h_t, h_t]$ can be obtained.

After the training of BiLSTM, the entity labeling of unlabeled words can be predicted from the output h_t . But in DNER task, some impossible combinations may also exist in the predicted data. For example, the label "I-BRAND" must not immediately follow the label "B-DRUG" logically, which means that we have to consider the label information of neighboring data. CRF is an undirected graphical model that focuses on the sentence level, instead of each position. Therefore, some impossible combinations should be ruled out.

With respect to the input sequence $Y = \{y^1, y^2, \dots, y^n\}$, y^n denotes the i th word vector of input, $Z = \{z^1, z^2, \dots, z^n\}$ is the label sequence of the input sequence Y , and P is the score matrix of output by BiLSTM, where k denotes score of the j th label of the i th word, and its score can be defined as follows:

$$s(Y, Z) = \sum_{i=0}^n A_{z_i, z_{i+1}} + \sum_{i=1}^n P_{i, z_{i+1}}, \tag{3}$$

where A is the transition score matrix, $A_{i, j}$ denotes the conversion score from the label i to the label j , and y_0 to y_n is

the start and end label of a sentence. They are added to a set of possible labels. Thus, A is a matrix whose size is $k+2$.

The loss function of CRF is composed of the actual path score and the total score of all possible paths; both scores are given as follows:

$$\begin{aligned} P_{\text{Realpath}} &= e^{s(Y,Z)}, \\ P_{\text{total}} &= \sum_{\tilde{z} \in Z_Y} e^{s(Y,\tilde{z})}, \end{aligned} \quad (4)$$

where $e^{s(Y,Z)}$ denotes the score of the possible path along, where the Z label is generated on the word Y and e is a numeric constant. In the training course, the log probability of the correct label sequence is maximized.

$$\log(P(Z|Y)) = \log\left(\frac{P_{\text{Realpath}}}{P_{\text{total}}}\right),$$

Lossfunction = $-\log(P(Z|Y))$,

$$= \left(\sum_{i=1}^L x_{iy_i} + \sum_{i=1}^{L-1} t_{y_i y_{i+1}} - \log\left(\sum_{\tilde{z} \in Z_Y} e^{s(Y,\tilde{z})} \right) \right). \quad (5)$$

The loss function of CRF is computed by formula (5), where x_{iy_i} denotes the emission score with the word index as i and the label index as y_i and $t_{y_i y_{i+1}}$ denotes the transmit score with the word index as y_i and the label index as y_{i+1} . Then, we can search for the optimal path using the Markov hypothesis, coupled with the Viterbi algorithm.

3.3. Multitask Learning Strategy. Multitask Learning (MTL) is a kind of joint learning through which the differences and connections between tasks can be effectively analyzed and modeled. Hard sharing, soft sharing, and hierarchical sharing are currently the most-used structures by MLT. Hard sharing stacks a given task on top of the sharing layer [39]. Soft sharing supports each task with separate models and parameters, and the internal information contained in each model can be accessed [40], but it may also lead to the inefficiency of parameters. Hierarchical sharing puts different tasks in different network layers [41], but it relies on the handmade hierarchical shared structure. For DNER, since the same entity has a number of synonyms and various forms of representations, exact matching or fuzzy matching as lookup methods of the dictionary may cause great challenges to detecting entity boundaries. However, this can be avoided by adding the DNEN task. Specifically, the output of DNER such as ‘‘B-DRUG’’ is an explicit signal indicating the start of drug entity so that the search space of DNEN can be reduced, vice versa. Therefore, two explicit feedback strategies were incorporated as a part of the multitask learning framework to simulate the reciprocal enhancement effect among different tasks.

A multitask learning framework resembling that proposed by Zhao [42] was used to enable DNER and DNEN to support each other and to enhance the generalization ability

of the model. In the first step, the training set was divided into subsets applicable to T tasks: D_1, \dots, D_T prior to the training process. In the training process, a training set t was chosen and the instance for random training $(w_{1:n}, y_{1:n}^t) \in D_t$ was acquired, where $w_i \in W$ and W denotes the input set; $y_i^t \in L^t$ and L^t denotes the label set. The label specific to the task t was used to predict the label y_i^t and update the label y_i^t and then the updated parameters were backed to the model for asynchronous training of DNER and DNEN, with the particular equation written as shown in Figure 4. where $\text{DNER}(w_{1:n}, i)$ and $\text{DNEN}(w_{1:n}, i)$ denote the DNER and the DNE normalized function with the word sequence w_1, w_2, \dots, w_n and the index i as inputs, y_{DNER}^i is the output of entity recognition applicable to the named entity label, y_{DNEN}^i is the output of the entity normalized function applicable to the entity vocabulary label, v_i^{DNER} is the input of DNER multiclass classification function that denotes the input of BiLSTM-CNN and the explicit feedback of DNEN, v_i^{DNEN} is the input of DNEN multiclass classification function that denotes the input of BiLSTM-CNN and the explicit feedback of DNEN. U is the matrix mapping from DNEN to DNER, and V is the matrix mapping from DNER to DNEN.

$$\begin{aligned} \text{DNER}(w_{1:n}, i) &= y_{\text{DNER}}^i = \arg \max y_{\text{DNER}}^i = f_{\text{DNER}}(v_i^{\text{DNER}}), \\ \text{DNEN}(w_{1:n}, i) &= y_{\text{DNEN}}^i = \arg \max y_{\text{DNEN}}^i = f_{\text{DNEN}}(v_i^{\text{DNEN}}), \\ v_i^{\text{DNER}} &= v^k \circ (v^k + y_{\text{DNEN}}^i U), \\ v_i^{\text{DNEN}} &= v^k \circ (v^k + y_{\text{DNER}}^i U), \\ F_{\theta}^k(x_{1:n}, i) &= v_i^k = h_{L,i}^k \circ h_{R,i}^k. \end{aligned} \quad (6)$$

In this article, a fully shared mode was adopted to make the BiLSTM-CNN layer shared among tasks, which means that all parameters as contained in the model would be shared, except for the output layer applicable to DNER and DNEN. This construction enables the proposed model to capture feature representations of different tasks and interactively give feedback to generate prediction sequences.

4. Network Training

In this section, we provided particular information in relation to training neural networks, including corpus, hyperparameter, optimizer, and assessment criteria. PyTorch was used to deploy the model and run the proposed model on Nvidia GTX 1080.

4.1. Datasets and Preprocessing. Obtain data from the DDI2011 and DDI2013 challenge corpora to construct the data set for training the deep learning model, and preprocess the data set for training the deep learning model in the following ways: randomly divide the dataset into T subsets, and T is an integer greater than or equal to 2. Establish four alphabets of word, character char, label label, and feature for each subset. Each alphabet is a dictionary for storing {key: instance, value: index}, where key represents the stored key,

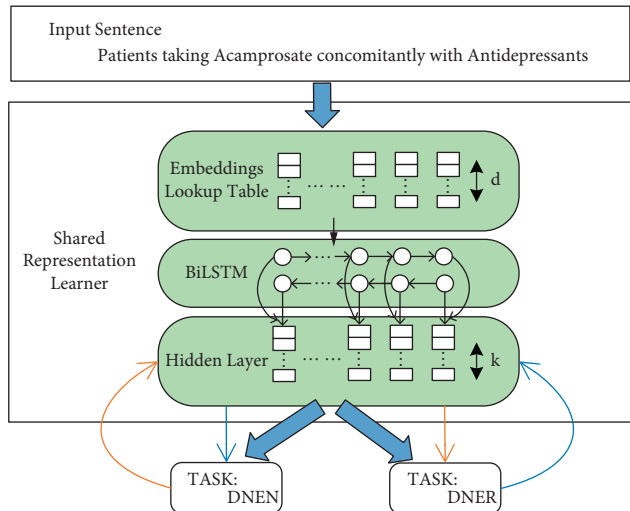


FIGURE 4: The main architecture of the multi-DTR model.

value represents the stored value, instance refers to the word, and index refers to the index. Based on the four alphabets of each subset, two lists are established for each subset. The two lists contain four columns of data, respectively. The four columns of data in the first list are [words, chars, labels, features], and the four columns of data in the second list are [words_Ids, chars_Ids, labels_Ids, features_Ids].

In the experiment, the DDI2011 Challenge Corpus from the drug-medicine interaction task was used. The minidom module as a part of python was used to extract <sentence> and <entity> elements, get the essential text and entity information, create a list, and match and annotate the entity and text. Next, all training datasets were collected as training data, and all test datasets were collected as test data. In this work, the sample was preprocessed using BIO labeling, where B denotes the first token of the entities in the sample, I denotes the token in the entity, and O denotes the token that does not fall into the category of entities. Table 1 lists the distribution of documents, sentences, and drugs as contained in the training and test set of DDI2011 [6]. Since there is only one type of entity names (DRUG) in this corpus, the text would be only labeled as “B/I-DRUG” or “O”.

For further performance evaluation of the proposed model, the SemEval-2013 dataset in drug name recognition and classification task was used. Table 2 shows the numbers assigned to the annotated entities in DDI2013 training set and test set. The dataset contains four entity types: Drug, Brand, Group, and Drug_n [43]. Drug denotes any chemical reagent served to treat, cure, prevent or diagnose human diseases. Brand is characterized by trade name or brand name. Group denotes any term that specifies the chemical or pharmacological relations between a group of drugs as mentioned in the text, and Drug_n describes a kind of chemical reagent that has not been approved for human medical use.

4.2. Pretrained Embedding. In this work, Pennington et al. [25] was used to initialize the word embedding obtained

TABLE 1: Training and testing set in DDI2011.

Set	Documents	Sentences	Drugs
Training	435	4267	11260
Final test	144	1539	3689
Total	579	5806	14949

from the pretraining on PMC and PubMed, and the context-sensitive word vectors were acquired using ELMo. The character embedding was randomly initialized according to a uniform sample $[-\sqrt{(3/\text{dim})}, +\sqrt{(3/\text{dim})}]$, where $\text{dim} = 30$.

4.3. Hyperparameters. Table 3 lists the hyperparameters used in the course of experiment. The dimensions of pre-trained word embedding, character embedding, and contextualized character embedding were set to 30, 100, and 1024, respectively. In the training process, the parameters were updated using Minibatch Stochastic Gradient Descent (SGD) in respect of descending learning rate. The initial learning rates of the proposed model, Dropout rate, and the batch size were set to 0.015, 0.5, and 10, respectively.

4.4. Criteria for Evaluation. In the experiment, the system performance was evaluated by precision, recall rate, and $F1$. Precision represents all correctly predicted entities as a percentage of all predicted entities. Recall rate represents the predicted entities as a percentage of all entities as contained in the dataset. $F1$ represents the harmonized mean value of precision and recall rate, with the following equation:

$$P = \frac{TP}{TP + FP},$$

$$R = \frac{TP}{TP + FN}, \quad (7)$$

$$F1 = \frac{2 * P * R}{P + R},$$

where TP denotes the number of true-positive samples, TN denotes the number of true-negative samples, FP denotes the number of false-positive samples, and FN denotes the number of false-negative samples. Two out of four criteria for evaluation available in DDI2013 [43] Challenge Corpus were used: type matching (only if there are some overlaps with the same category of gold drug names) and strict matching (only if the label boundary and category are the same as the gold drug names, the label drug names are correct).

5. Experiment and Analysis

The multi-DTR model as described here was evaluated on DDI2011 and DDI2013, known as the representative biomedical corpora. Table 4 is the performance comparison of multi-DTR with the works done by other teams. Next, the impact of each architecture (e.g., different embedded layers, different optimization methods, and multitask mutual feedback framework) as a part of the proposed model on the

TABLE 2: Numbers of the annotated entities in DDI2013 set.

Type	Train			Test		
	DrugBank	MedLine	Total	DrugBank	MedLine	Total
Drug	9901 (63%)	1745 (63%)	11646 (63%)	180 (59%)	171 (44%)	351 (51%)
Brand	1824 (12%)	42 (1.5%)	1866 (10%)	53 (18%)	6 (2%)	59 (8%)
Group	3901 (25%)	324 (12%)	4225 (23%)	65 (21%)	90 (24%)	155 (23%)
Drug_n	130 (1%)	635 (23%)	765 (4%)	6 (2%)	115 (30%)	121 (18%)
Total	15756	2746	18502	304	382	686

TABLE 3: The parameters for our experiments.

Layer	Hyperparameter	Value
CNN	Window size	3
	Number of filters	30
LSTM	State size	200
	Initial state	0.0
	Peepholes	No
Dropout	Dropout rate	0.5
	Batch size	10
	Initial learning rate	0.015
	Gradient clipping	5.0
	Decay rate	0.05
	Labeling schema	BIO
	ELMo dim	1024

experiment was assessed. The findings of comparison suggest that the architectures of the proposed model would perform well in the experiment.

5.1. Performance Comparison with Available Methods.

The results were compared with those of the works done by other teams. For the sake of fairness and rationality of the experiment, the hyperparameters of the proposed model were configured according to the optimal parameters as referred to in the article. As is seen from Table 4, the dictionary-based method and the rule-based method, as proposed earlier, yielded reasonable results, including Tsuruoka [12] and Hettne et al. [13], subsequent deep learning model. For example, LASIGE et al. [43] combined CRF with the list of dictionary terms intended for DNER processing as collected from the database in order to recognize and classify entities. Zeng et al. [16] used the BiLSTM-CRF structure to identify drug entities without the aid of any external dictionary, with good results attained. Yang et al. [37] used a hierarchical recursive network for cross-language transfer learning. The model proposed by Liu et al. [44] combines the word embedding trained in biomedical text with the semantic features of three drug dictionaries, with an impressive performance on DDI2013, suggesting that the accuracy of our proposed model is 0.90% lower than that proposed by Liu et al. [44], but its recall rate and F1 are 6.23% and 2.43% higher than that proposed by Liu [44].

For the evaluation of DDI2013 dataset, Table 5 provides a summary of the accurate evaluation of the proposed model in the entity type-specific recognition as part of DDI2013.

Despite good performance in type recognition, the proposed model may neglect the difference between a given

entity and other entity types due to a small percentage (<4%) of Drug_n entity type in the dataset. As a result, the recognition accuracy of the proposed model would be lower than that of any other entity.

5.2. Performance Comparison of Different Statements.

This work proposed using pretrained word embedding, character representation, and context-sensitive word embedding to obtain additional feature information, as given in Table 6. To test the impact of different input information representations on the proposed model, three kinds of embedding information were combined and input into the model, respectively. According to the results, serial representation is better than single representation, and multiple representations can attain the best performance.

5.3. Comparison of Optimization Methods.

Different optimizers, including SGD, AdaGrad, Adadelta, RMSProp, and Adam, were compared here. SGD can calculate gradient and update parameters by randomly extracting the training sample of a fixed size while avoiding falling into saddle points or poor local optimal points. AdaGrad imposes a constraint on the optimal learning rate and is suitable for processing sparse gradient, but it may cause the disappearance of gradient. Adadelta is an extension of AdaGrad and simplifies the computational process. RMSProp relies on a global learning rate and is suitable for processing non-stationary targets. Adam can adjust the parameter-specific learning rate using first-order moment estimation and second-order moment estimation, but it is vulnerable to generalization and convergence problems. According to the experimental results, as given in Figure 5, SGD is significantly better than any other optimizer.

5.4. Performance Comparison in Case of Dropout.

The effectiveness of Dropout was evaluated here, with all of the other hyperparameters in the model identical to that in Table 3. As given in Table 7, the performance of the proposed model on DDI2011 and DDI2013 was slightly improved after the Dropout was used, which in turn proves that Dropout plays a part in reducing overfitting.

5.5. Performance Comparison between Multitask Learning and Single-Task Learning.

The effectiveness of multitask learning strategy was also examined. As seen from Table 8, the efforts to jointly model DNER and DNEN by using two explicit feedback strategies would significantly improve the

TABLE 4: Results of experiment in DDI2011 and DDI2013.

System	DDI2011			DDI2013		
	Precision	Recall	F1	Precision	Recall	F1
UMCC_DLS	-	-	-	24.00	57.00	34.00
Hettne	66.91	71.42	69.09	59.41	56.32	57.82
Tsuruoka	68.42	72.39	70.34	62.24	58.17	60.12
WBI	89.53	88.42	88.97	76.70	88.42	74.80
LASIGE	87.02	82.51	84.70	78.00	56.00	65.19
Yang	81.44	81.50	81.46	76.54	74.40	75.45
Zeng	93.26	91.11	92.17	83.60	77.81	79.26
Liu	-	-	-	87.46	75.22	80.88
Multi-DTR	94.36	92.13	93.22	85.56	81.45	83.45

TABLE 5: Experimental results of different entity types in DDI2013.

Type	Precision	Recall	F1
Drug	86.52	81.68	84.03
Brand	89.46	78.51	83.62
Group	83.26	86.43	84.81
Drug_n	79.74	67.36	73.02
Mico-average	85.56	81.45	83.45

TABLE 6: Performance comparison of each representations.

System	DDI2011			DDI2013		
	Precision	Recall	F1	Precision	Recall	F1
ELMo	88.46	87.74	88.09	82.14	79.24	81.68
Char	86.32	85.12	85.71	81.21	78.53	79.84
Glove	88.12	89.34	88.72	84.74	80.57	82.60
ELMo + Char	89.47	90.55	90.00	83.45	81.06	82.23
Char + Glove	90.14	88.42	89.24	83.32	80.64	81.95
ELMo + Glove	91.73	89.51	90.60	84.24	81.32	82.75
ELMo + Glove + Char	94.36	92.13	93.22	85.56	81.45	83.45

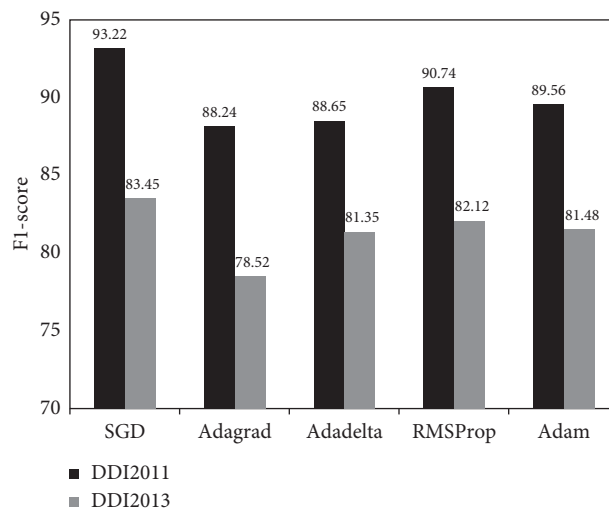


FIGURE 5: Performance comparison of different optimization methods optimization.

TABLE 7: Performance comparison using Dropout.

		Precision	Recall	F1
DDI2011	No	92.73	91.11	91.91
	Yes	94.36	92.13	93.22
	Δ	+1.63	+1.02	+1.31
DDI2013	No	83.52	79.71	81.04
	Yes	85.56	81.45	83.45
	Δ	+2.04	+1.74	+2.41

TABLE 8: Performance comparison of adopting multitask learning.

		Precision	Recall	F1
DDI2011	Single-task	91.13	89.51	90.30
	Multitask	94.36	92.13	93.22
	Δ	+3.23	+2.62	+2.92
DDI2013	Single-task	83.42	78.01	80.62
	Multitask	85.56	81.45	83.45
	Δ	+2.14	+3.44	+2.83

model performance, partly because the multitask learning provides a general representation of both tasks and partly because the proposed method converts hierarchical tasks into parallel multitask setting and retains mutual support between different tasks.

6. Conclusion

Drug text mining is a key interdisciplinary field of computer science and biomedicine. In this work, a multitask learning framework was tailored for DNER, with an impressive performance on DDI2011 and DDI2013. Through detailed analysis, the main gains of the proposed model can be attributed to character sharing between drug entities, pre-trained word embedding, and context-sensitive word embedding information. The conflict of entity boundary and type can be generally resolved by the positive feedback of DNER and DNEN. According to the experimental results, the proposed method can readily perform well without the aid of any drug dictionary or manual creation so an efficient DNER system was constructed.

Data Availability

The experimental datasets used in this work are publicly available, and the bundled data and code of this work 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 National Natural Science Foundation of China under Grant nos. 61472348 and 61672455, by the Humanities and Social Science Fund of the Ministry of Education of China under Grant no. 17YJCZH076, by Zhejiang Science and Technology Project

under Grant nos. LGF18F020001 and LGF21F020022, and by the Ningbo Natural Science Foundation under Grant no. 202003N4324.

References

- [1] U.S. Food and Drug Administration, "How FDA reviews proposed drug names," <http://www.fda.gov/downloads/Drugs/DrugSafety/MedicationErrors/ucm080867.pdf>.
- [2] S. Doan and H. Xu, "Recognizing medication related entities in hospital discharge summaries using support vector machine," in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters COLING '10*, pp. 259–266, Association for Computational Linguistics, Stroudsburg, PA, USA, 2010.
- [3] D. Sanchez-Cisneros and F. Aparicio Gali, "UEM-UC3M: an ontology-based named entity recognition system for biomedical texts," in *Second Joint Conference on Lexical and Computational Semantics (SEM)*, S. Manandhar and D. Yuret, Eds., vol. 2, Association for Computational Linguistics, Atlanta, Georgia, USA, 2013.
- [4] J. Björne, S. Kaewphan, and T. Salakoski, "UTurku: drug named entity recognition and drug-drug interaction extraction using SVM classification and domain knowledge," in *Second Joint Conference on Lexical and Computational Semantics (SEM)*, S. Manandhar and D. Yuret, Eds., vol. 2, Association for Computational Linguistics, Atlanta, Georgia, USA, 2013.
- [5] I. Segura-Bedmar, P. Martínez, and M. Herrero-Zazo, "Lessons learnt from the DDIExtraction-2013 shared task," *Journal of Biomedical Informatics*, vol. 51, pp. 152–164, 2014.
- [6] I. Segura-Bedmar, P. Martinez, and D. Sanchez-Cisneros, "The 1st DDIExtraction-2011 challenge task: extraction of drug-drug interactions from biomedical texts," in *Proceedings of the 1st Challenge Task on Drug-Drug Interaction Extraction*, pp. 1–9, Huelva, Spain, September 2011.
- [7] D. Rebholz-Schuhmann, A. Yepes, E. M. van Mulligen et al., "The CALBC silver standard corpus-harmonizing multiple semantic annotations in a large biomedical corpus," *Proceedings of the Third International Symposium on Languages in Biology and Medicine*, vol. 2009, pp. 64–72, 2009.
- [8] A. Ekbal and S. Saha, "Simultaneous feature and parameter selection using multiobjective optimization: application to named entity recognition," *International Journal of Machine Learning and Cybernetics*, vol. 7, no. 4, pp. 597–561, 2016.
- [9] O. Mai and S. Khaled, "Nera 2.0: improving coverage and performance of rule-based named entity recognition for Arabic," *Natural Language Engineering*, vol. 23, no. 3, pp. 441–447, 2016.
- [10] U. Kanimozhi and D. Manjula, "A CRF based machine learning approach for biomedical named entity recognition," in *Proceedings of the 2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)*, vol. 335–342, Tindivanam, India, February 2017.
- [11] S. Ananiadou and J. McNaught, *Text Mining for Biology and Biomedicine*, Artech House, Inc., Norwood, MA, USA, 2005.
- [12] Y. Tsuruoka, J. McNaught, J. Tsujii, and S. Ananiadou, "Learning string similarity measures for gene/protein name dictionary look-up using logistic regression," *Bioinformatics*, vol. 23, no. 20, pp. 2768–2774, 2007.
- [13] K. M. Hettne, R. H. Stierum, M. J. Schuemie et al., "A dictionary to iDNENtify small molecules and drugs in free text," *Bioinformatics*, vol. 25, pp. 2983–2991, 2009.

- [14] R. Eriksson, P. B. Jensen, S. Frankild, L. J. Jensen, and S. Brunak, "Dictionary construction and identification of possible adverse drug events in Danish clinical narrative text," *Journal of the American Medical Informatics Association*, vol. 20, pp. 947–953, 2013.
- [15] A. Cocos, A. G. Fiks, and A. J. Masino, "Deep learning for pharmacovigilance: recurrent neural network architectures for labeling adverse drug reactions in Twitter posts," *Journal of the American Medical Informatics Association*, vol. 24, no. 4, pp. 813–821, 2017.
- [16] D. Zeng, C. Sun, L. Lin, and B. Liu, "LSTM-CRF for drug-named entity recognition," *Entropy*, vol. 19, no. 6, p. 283, 2017.
- [17] J. Devlin, M.-W. Chang, K. Lee, and T. Kristina, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," <https://arxiv.org/abs/1810.04805>.
- [18] J. Lee, W. Yoon, S. Kim et al., "BioBERT: A Pre-trained Biomedical Language Representation Model for Biomedical Text Mining," 2019, <https://arxiv.org/abs/1901.08746>.
- [19] S. Pradhan, N. Elhadad, W. Chapman, S. Manandhar, and G. Savova, "Semeval-2014 Task 7: Analysis of Clinical Text," in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pp. 54–62, Dublin, Ireland, January 2014.
- [20] C. H. Wei, Y. Peng, R. Leaman et al., "Overview of the biocreative v chemical disease relation (CDR) task," in *Proceedings of the Fifth BioCreative Challenge Evaluation Workshop*, Seville, Spain, September 2015.
- [21] K. Ning, B. Singh, Z. Afzal, M. Erik M van, and A. K. Jan, "Using rule-based natural language processing to improve disease normalization in biomedical text," *Journal of the American Medical Informatics Association*, vol. 20, no. 5, pp. 876–881, 2013.
- [22] H. C. Lee, Y. Y. Hsu, and H. Y. Kao, "Audis: An Automatic Crf-enhanced Disease Normalization in Biomedical Text," *Database 2016*, vol. 2016, 2016.
- [23] Y. Lou, Y. Zhang, T. Qian, L. Fei, X. Shufeng, and J. Donghong, "A transition-based joint model for disease named entity recognition and normalization," *Bioinformatics*, vol. 33, no. 15, pp. 2363–2368.
- [24] T. Mikolov, K. Chen, G. Corrado, and D. Jeffrey, "Efficient estimation of word representations in vector space," 2013, <https://arxiv.org/abs/1301.3781>.
- [25] J. Pennington, R. Socher, and C. Manning, "Glove: global vectors for word representation," in *Proceedings of the Conference on Empirical Methods in Natural Language*, pp. 65–72, Doha, Qatar, January 2014.
- [26] T. Yu and R. Jin, "Research review of natural language processing pretraining model," *Computer engineering and application*, vol. 56, no. 23, pp. 12–22, 2020.
- [27] X. Ma and E. Hovy, "End-to-end sequence labeling via bidirectional lstm-cnns-crf," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp. 1064–1074, Berlin, Germany, March 2016.
- [28] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuska, "Natural language processing (almost) from scratch," *Journal of Machine Learning Research*, vol. 12, pp. 2493–2537, 2011.
- [29] L. Wang, C. Dyer, A. W. Black et al., "Finding function in form: compositional character models for open vocabulary word representation," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 1520–1530, Lisbon, Portugal, August 2015.
- [30] C. D. Santos and B. Zadrozny, "Learning character-level representations for part-of-speech tagging," in *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pp. 1818–1826, Beijing, China, July 2014.
- [31] O. Kuru, O. A. Can, and D. Yuret, "Charner: Character-Level Named Entity Recognition," in *Proceedings of the 26th International Conference on Computational Linguistics*, pp. 911–921, Osaka, Japan, 2016.
- [32] Q. Tran, A. MacKinlay, and A. J. Yepes, "Named entity recognition with stack residual LSTM and trainable bias decoding," pp. 566–575, 2017, <https://arxiv.org/abs/1706.07598>.
- [33] Z. Zhao and Y. Wu, "Attention-based convolutional neural networks for sentence classification," in *Proceedings of the 17th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, pp. 705–709, San Francisco, CA, USA, September 2016.
- [34] E. Strubell, P. Verga, D. Belanger, and A. McCallum, "Fast Andaccurate Entity Recognition with Iterated Dilated Convolutions," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 2670–2680, ACL, January 2017.
- [35] J. P. C. chiu and E. Nichols, "Named En-Tity Recognition with Bidirectional LSTM-CNN s," 2015, <https://arxiv.org/abs/1511.08308>.
- [36] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult. Neural Networks," *IEEE Transactions on*, vol. 5, no. 2, pp. 157–166, 1994.
- [37] Z. Yang, R. Salakhutdinov, and W. W. Cohen, "Transfer learning for sequence tagging with hierarchical recurrent networks," in *Proceedings of the 5th International Conference on Learning Representations*, pp. 1–10, Toulon, France, 2017.
- [38] Z. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," 2015, <https://arxiv.org/abs/1508.01991>.
- [39] R. Collobert and J. Weston, "A unified architecture for natural language processing: deep neural networks with multitask learning," *Proceedings of the 25th International Conference on Machine Learning*, pp. 160–167, Helsinki, Finland, 2008.
- [40] I. Misra, A. Shrivastava, A. Gupta, and H. Martial, "Cross-stitch networks for multi-task learning," in *Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3994–4003, Las Vegas, NV, USA, 2016.
- [41] N. Srivastava, G. E. Hinton, and A. Krizhevsky, "Dropout: a simple way to prevent neural networks from over fitting," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2016.
- [42] S. Zhao, T. Liu, S. Zhao, and W. Fei, "A neural multi-task learning framework to jointly model medical named entity recognition and normalization," *AAAI*, vol. 33, pp. 817–824, 2019.
- [43] I. Segura-Bedmar, P. Martínez, and M. Herrero-Zazo, "Lessons learnt from the DDI Extraction-2013 shared task-Science Direct," *Journal of Biomedical Informatics*, vol. 51, pp. 152–164, 2014.
- [44] S. Liu, B. Tang, Q. Chen, and X. Wang, "Effects of semantic features on machine learning-based drug name recognition systems: word embeddings vs. manually constructed dictionaries," *Information*, vol. 6, pp. 848–865, 2015.

Research Article

Prediction of Purchase Volume of Cross-Border e-Commerce Platform Based on BP Neural Network

Xiang Zhang 

Zhuhai College of Science and Technology, Zhuhai, Guangdong 519040, China

Correspondence should be addressed to Xiang Zhang; 13411393334zx@z cst.edu.cn

Received 1 March 2022; Revised 21 March 2022; Accepted 24 March 2022; Published 15 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiang Zhang. 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.

As a new form of foreign trade, cross-border e-commerce has huge development potential. Although the development prospect of cross-border e-commerce is good, the management of global supply chain is very important in order to gain a place in the fierce competition and develop steadily. The traditional forecasting of purchasing volume adopts time series, and the forecasting model is relatively simple. The purchase volume of the platform is related to the various consumption behaviors of consumers, such as the number of product reviews, the number of product collections, and whether there are tax subsidies. The sales volume in the next few days is predicted by the item number, time, and sales quantity. The four-layer BP neural network model is used, and the MATLAB neural network toolbox is used to draw the training error curve and the correlation coefficient curve. After network training, the training correlation coefficient R reaches 95.823%, and the prediction accuracy obtained at this time is higher. Further, using the established model based on BP algorithm, the traditional BP algorithm is optimized to obtain the purchase quantity of commodities. The method is applied to the forecast of commodity purchase volume of a cross-border e-commerce platform, and the results show that the average error rate of this method is 5.9%, which has high practical application value. The research results show that this paper considers multiple influencing factors and selects an appropriate forecasting method, which can effectively improve the accuracy of the company's commodity sales forecast, so as to better formulate procurement plans and optimize inventory structure, which has certain implications for the actual operation of cross-border e-commerce platforms.

1. Introduction

Today, with the development of information technology, people have gradually entered the information age, and traditional product sales methods have been unable to meet people's shopping needs. In the e-commerce work, no matter the size of the seller's store, the problem that the seller must be unable to avoid is the replenishment of orders, and the two main problems faced by the replenishment of orders are the occupation of cash flow and the problem of inventory. If the replenishment order is not predicted well, the capital occupation will have a very large impact on the capital operation of the merchant, and the backlog of inventory will also increase the storage cost and even cause the capital flow to slow down. In addition, this paper selects cross-border e-commerce platforms (Figure 1) to model and predict sales, which further narrows the scope of the study.

This idea provides a new perspective for studying cross-border e-commerce platform purchases. In addition, when selecting influencing factors, variable indicators related to the platform itself are introduced, which enrich the types of factors currently used to study the purchase volume of cross-border e-commerce platforms. By comparing the forecasting results of different models, a sales forecasting model with higher forecasting accuracy that is different from the traditional sales forecasting method is determined. On the other hand, by constructing an appropriate sales volume forecast model and forecasting sales volume, the e-commerce platform can accurately formulate policies and marketing strategies that are consistent with the current economic environment and the development of the e-commerce market according to the forecast results and promote the e-commerce platform: healthy and stable development. At the same time, it also helps the company's

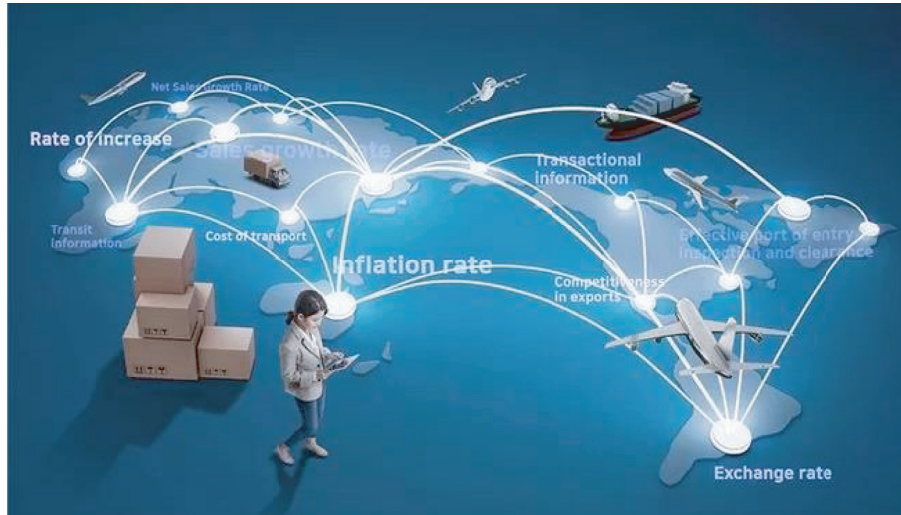


FIGURE 1: Cross-border e-commerce.

regulatory department to monitor the company's production capacity in real time, arrange production plans reasonably, avoid the problem of excess production capacity, and help the company to reduce inventory costs and reduce marketing pressure. Therefore, it is of positive significance to accurately predict the purchase volume of cross-border e-commerce platforms [1–5].

Qualitative forecasting method, also known as empirical judgment method, is mainly used by forecasters to estimate and judge the future development trend of the market environment based on personal knowledge and experience, plus various information obtained through market research. Qualitative prediction method has obvious advantages for start-ups, it is time consuming, simple, and easy to operate and can comprehensively consider a variety of factors. The disadvantage is that it is subjective and arbitrary, and the forecast results are not accurate enough. It is more suitable for long-term market environment forecasting than short-term forecasting. Commonly used qualitative prediction methods include comprehensive opinion method, subjective probability method, expert evaluation method, and Delphi method. Qualitative prediction does not need to analyze a large amount of data and does not require relevant mathematical statistics. It only relies on professionals' knowledge of the industry and their grasp of the actual situation to make judgments [6–10].

Compared with qualitative forecasting, quantitative forecasting abandons most of the subjective factors and no longer highly depends on the individual ability of the forecaster. At present, quantitative forecasting methods are more widely used. Quantitative prediction excludes the subjective experience of people, and to a greater extent, on the basis of facts and data, it is necessary to find more historical data, and use computers to establish data models to explore the potential laws in the data to calculate the prediction results. In short-term forecasting, quantitative forecasting is more accurate than qualitative forecasting, but it requires certain professional skills for forecasters. At present, quantitative forecasting is mainly divided into time

series forecasting method, causality forecasting method, and other forecasting forecasting methods [11–15].

Data mining forecasting mainly includes the following three parts: algorithms and techniques, data, and modeling. Among them, the selection and optimization of prediction algorithms and techniques are the top priorities of experts' research in recent years. The research focus of this paper is based on the problem of forecasting the purchase volume of cross-border e-commerce platforms. Through comparison, the algorithm model with the best prediction effect is selected to complete the forecast of cross-border e-commerce platform purchase volume. First, a suitable prediction algorithm needs to be determined. As an important link in the data mining forecasting process, the selection of the forecasting algorithm also prepares for the subsequent accurate forecast of the purchase volume of the cross-border e-commerce platform. Machine learning prediction algorithm is a commonly used prediction algorithm in data mining prediction analysis. The following will focus on the basic principles and advantages and disadvantages of several machine learning prediction algorithms and choose the most suitable algorithm to complete the prediction after optimization [16–19].

Machine learning is a technology that has emerged in recent years and has been widely used in many fields, such as data mining, computer vision, natural language processing, biometric recognition, search engines, medical diagnosis, stock market analysis, and DNA sequence testing. Assuming that there is a functional relationship $y = f(xB)$ between the input and output of things, where y is the undetermined parameter and x is the input variable, then $y = f(xB)$ is called a learning machine. Through modeling, the model is repeatedly trained by the sample data to obtain the value of parameter B , and the functional relationship between x and y in $y = f(xB)$ is determined, so that the y value of the new x can be predicted. This process is a typical machine learning process. Among them, the classic machine learning prediction methods include decision tree method, support vector machine, Bayesian network, time series model, neural network, and so on [20–22].

There are many kinds of forecasting methods commonly used at present, mainly including time series forecasting method, smoothing forecasting method, trend line forecasting method, seasonal cycle forecasting method, regression forecasting technique, and so on. Although these methods have been applied in various occasions, they have the following deficiencies more or less: (1) most of the models can only be applied to linear situations and cannot be applied well to nonlinear situations; (2) it is difficult to apply to multifactor situations; and (3) the establishment of the model depends on the forecaster's understanding of specific problems and experience in forecasting.

Artificial neural network is a nonlinear system that simulates the information processing method of the human brain. It has the characteristics of distributed storage and parallel processing of knowledge and has the functions of memory and association. Theoretical research shows that it can approximate any nonlinear function with arbitrary precision, so it is more suitable for modeling some complex problems. Prediction of sales volume is undoubtedly very important for operators, but there are many factors that affect sales volume. Under the current situation in my country, it is also related to many national policies, etc., and it is difficult to accurately predict with conventional methods. To this end, this paper discusses the use of neural networks to establish a sales forecast model, uses the memory ability of neural networks to learn and memorize historical data, establishes a neural network model of sales, and then uses the associative ability of neural networks to predict future sales [23–25].

2. BP Neural Network

BP neural network is a multilayer feed-forward network, which has been developed so far and is one of the most widely used neural network models in various industries. Its network topology consists of an input layer, several hidden layers, and an output layer. Among them, the three-layer network structure diagram is shown in Figure 2. BP neural network has complex classification ability and strong nonlinear modeling and mapping ability and is widely used in complex prediction problems. The BP neural network can continuously repair the error between the expectation and the sample data, which is derived from its calculation and adjustment of its own weight threshold to make corrections. However, this model also has certain problems. The BP algorithm has a slow convergence speed, and the BP algorithm is easy to fall into local minimization.

There are many types of neural network structures. This paper adopts a commonly used forward neural network model to solve the problem of sales forecasting. The most commonly used learning algorithm for the forward neural network is the BP algorithm, but the BP algorithm is easy to fall into the local minimum. Evolutionary algorithm is an optimization method based on group optimization. Its basic idea is to simulate the evolution process of biological groups in nature. According to Darwin's evolutionary thought, in nature, biological groups compete with each other and proceed according to the principle of survival of the fittest.

Selection of individuals who adapt to the environment is retained, and individuals who do not adapt to the environment are eliminated. Individuals can inherit the characteristics of their parents, and at the same time mutate to have different characteristics from their parents, which provides more and different individuals for selection in nature. This is how the group evolves from generation to generation, and finally evolves from being incompatible with the environment to adapting to the environment. Generally, there are three different types of evolutionary algorithms: genetic algorithms, evolutionary programming, and evolutionary strategies. Due to the characteristics of neural network learning, the learning of evolutionary strategies has certain advantages. The specific steps of the BP algorithm can be simply summarized as follows:

- (1) Given the input vector X and target output vector T of the network and initialize the neural network weights;
- (2) Calculate the actual output of the network;
- (3) Calculate the error between the actual output vector of the network and the required target output value;
- (4) Weight learning to minimize the error;

Then repeat steps (1)–(4) to minimize the systematic error.

In the training process of BP neural network, in order to reduce the error between the expected value and the actual value, each connection weight is corrected layer by layer according to the reverse direction of the forward propagation process. It is called the error back-propagation algorithm. With the continuous correction of error back-propagation training, the correct rate of the network's response to the input mode will also continue to improve. The whole process includes forward propagation and back propagation. These two processes are described in detail below:

Step 1: forward propagation process. Let the weight between node i and node j be W_{ij} , the threshold of node j is b_j , and x is the output value of the node. Substitute w_{ij} and b_j into the activation function to obtain the output value of each node according to the following formula:

$$S_j = \sum_{i=0}^{m-1} w_{ij}x_i + b_j, \quad (1)$$

$$x_j = f(S_j),$$

where f is the activation function.

Step 2: backpropagation process. In the reverse direction of forward propagation, the output layer passes through the hidden layer, and finally returns to the input layer to correct the weights and thresholds of each neuron layer by layer until the error between the expected value and the predicted value reaches the minimum. Assuming that the result of the output layer is d_j and the expected output is y_j , the error function is as follows:

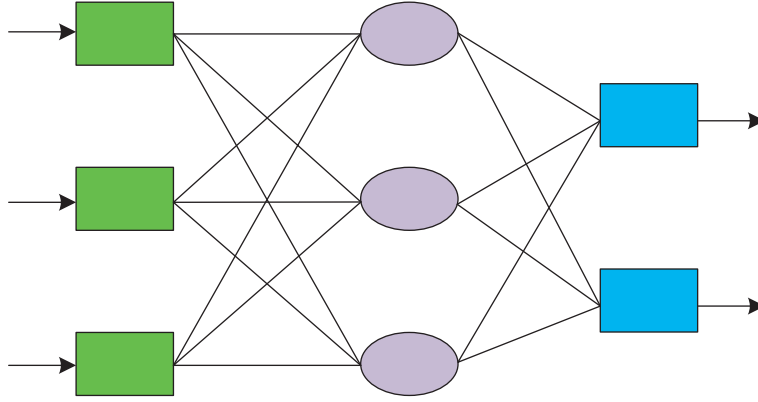


FIGURE 2: The three-layer network structure diagram.

$$E(w, b) = \frac{1}{2} \sum_{j=0}^{n-1} (d_j - y_j)^2. \quad (2)$$

It can be seen from the gradient descent method adopted by the traditional BP neural network that the correction value of the weight is proportional to the gradient of $E(w, b)$, then for the j th output node is

$$\Delta w(i, j) = \eta \frac{\partial E(w, b)}{\partial w(i, j)}, \quad (3)$$

where η is the learning rate. If the activation function is selected as

$$f(x) = \frac{A}{1 + e^{-x/B}}. \quad (4)$$

That is, the weights and thresholds between the hidden layer and the output layer are adjusted according to the following formula:

$$\begin{aligned} w_{ij} &= w_{ij} - \eta_1 \delta_{ij} x_i, \\ b_j &= b_j - \eta_2 \delta_{ij}. \end{aligned} \quad (5)$$

Similarly, according to the above method, the w_{ij} and b_j between the input layer and the hidden layer are adjusted as follows:

$$\begin{aligned} w_{ki} &= w_{ki} - \eta_1 \delta_{ki} x_k, \\ b_i &= b_i - \eta_2 \delta_{ki}. \end{aligned} \quad (6)$$

Among them, η is the learning rate, δ is the local gradient, and x_k is the output signal of the previous layer. The predicted value is shown in Figure 3.

Once the representation of the neural network is determined, the next step in evolutionary learning is to generate the neural network by randomly generating strings. The performance index of the neural network is related to the cumulative error sum of squares:

$$f = F_{\max} - \sum_{i=1}^N (y_i - \hat{y}_i). \quad (7)$$

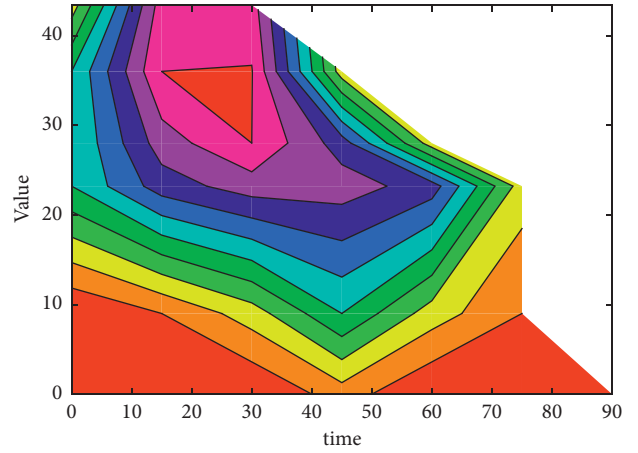


FIGURE 3: Predicted value.

In the formula, y_i is the expected output, which comes from the sample (x_i, y_i) ; y is the actual output of the neural network when the input is x_i ; and F_{\max} is a preset coefficient. The larger the f is, the closer the neural network is to the sample point, so its corresponding performance is better.

Different from the genetic algorithm, the evolution strategy does not use the crossover operator, it only uses the mutation operator, which can mutate the topology and weights of the neural network at the same time. For simplicity, this paper assumes that the topology of the neural network is fixed (using forward neural network, the number of neurons in each layer is fixed), so it is only necessary to mutate the weights. The weight mutation is to add an increment to the original weights. For the weight of the n th neural network in the group, the increment is

$$\begin{aligned} \Delta w_{ij}^k &= N(0, p_m), \\ p_m &= \lambda \left(\frac{1-f}{f_{\max}} \right). \end{aligned} \quad (8)$$

In the formula, λ is the coefficient, $N(0, p_m)$ is a Gaussian function with a mean of 0 and a variance of p_m , and p_m is related to the relative performance of the n th neural network

in the population, if the performance of the n th neural network is worse, the greater the pm , the greater the possibility that Δw deviates from the mean value of 0.

From the learning process of the evolution strategy to the neural network, it can be seen that the evolution strategy adopts the method of group learning, so it can avoid falling into the local extreme point because it only uses the mutation operator, if necessary, it can the topology of the neural network is learned. The predicted data are shown in Figure 4.

3. BP Neural Network Modeling

When analyzing the influencing factors of the purchase volume of cross-border e-commerce platforms, this paper selects the desensitization data set of food, cosmetics, and clothing from a self-operated cross-border e-commerce platform from January 2003 to September 2018. A total of 13 influencing factor indicators were selected. Through correlation analysis and gray correlation analysis, it was found that the selected factors and indicators had high correlation and correlation with the procurement of cross-border e-commerce platforms with a lag of one period. Therefore, when building the BP neural network model, refer to the analysis results, and still select the number of reviews, unit price, good rating, 7-day return, season, tax available, discount available, inventory, purchase rate, number of items added to the shopping cart, sales, and items. The neural network prediction model is constructed by 13 indicators such as type, number of advertisement-guided views, and so on. Since the measurement units of these influencing factors are different and the value ranges are very different, the data also need to be standardized before constructing the BP neural network model.

After determining the variables that affect the purchase volume of the cross-border e-commerce platform, divide the collected data into training data and test data. The division method is consistent with the construction of the linear regression model, where the output variable (Y) is the cross-border e-commerce lag period by one period. The input variable (X) is the 13 influencing factors that are highly correlated and correlated with the purchase volume of cross-border e-commerce platforms selected in this paper. Due to the small amount of data, 80% of the data is divided into training data and the remaining 20% is used as test data. The data of various influencing factors from 2017 to 2019 are compared with the purchase volume data of cross-border e-commerce platforms from 2017 to 2018. The combination is used as the data for the training model, and the data of each influencing factor from 2018 to 2019 and the purchase volume data of the cross-border e-commerce platform from 2017 to 2019 are combined as the data for the test model. After the data are divided, the BP neural network model is constructed by using the standardized data of each index. The details are as follows:

- (1) Determine the number of layers of the network. When dealing with the vast majority of specific problems, the network structure consisting of three levels can basically deal with it. Therefore, when

studying the purchase volume of cross-border e-commerce platforms in this paper, due to the small amount of data for each variable index, we choose to build a three-layer BP neural network model with only one hidden layer.

- (2) Enter the setting of the number of nodes in this layer. When getting a specific research question, the number of variables involved in the research object determines how many nodes to design in this layer. After analyzing the influencing factors of cross-border e-commerce platform procurement volume, this paper determines 13 influencing factors to model and analyze the cross-border e-commerce platform procurement volume, so this layer is designed as 13 nodes.
- (3) Output the setting of the number of nodes in this layer. For a practical research problem, when determining how many nodes to design in this layer, it should also be based on the actual situation. The problem studied in this paper is to predict the purchase volume of cross-border e-commerce platforms, so the number of nodes in this layer is set to 1.
- (4) Setting the number of hidden layer nodes. When constructing the network structure, the number of hidden layer nodes plays a very critical role. However, up to now, the research has found that there is still no ideal solution to the problem of how to select an appropriate number of hidden nodes for the hidden layer method, so this paper uses the empirical formula summarized by the researchers to determine the number of hidden layer nodes m .
- (5) Select the transfer function used in the network model. When constructing the network model in this paper, the Sigmoid function, which is widely used in the nonlinear transfer function, is selected because it has the characteristics that the linear function does not have, and its smoothness and differentiability make it occupy a special position in the transfer function. And it also has strong fault tolerance performance. The output layer directly selects a pure linear function.
- (6) Determine the method for training the network model and the parameters involved in the network model. In this paper, when using the training data to train the network model, a stochastic gradient descent (sgd) strategy, which is different from the standard BP algorithm, is selected. This algorithm randomly selects data with the same sample size for training, and then only uses this part of the sample data to calculate the gradient. In addition, set the number of epochs of model training (epochs) to 1000 and finally determine the learning rate of 0.09 and the momentum coefficient of 0.85 through multiple test adjustments.

Based on the results of the above-determined parameter values, the BP neural network model is finally constructed, and the network structure is shown in Figure 5.

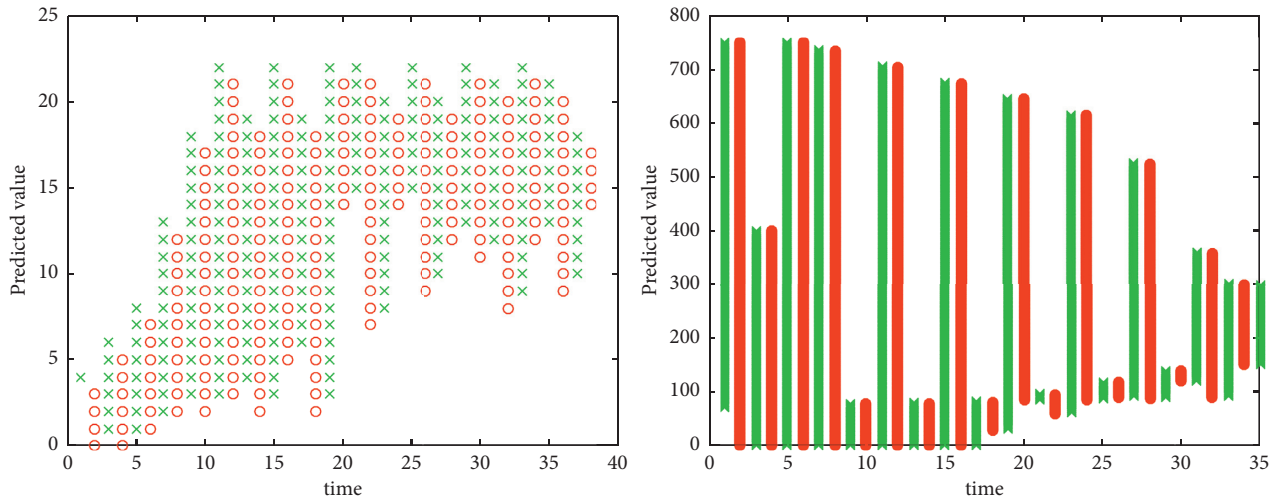


FIGURE 4: Predicted data.

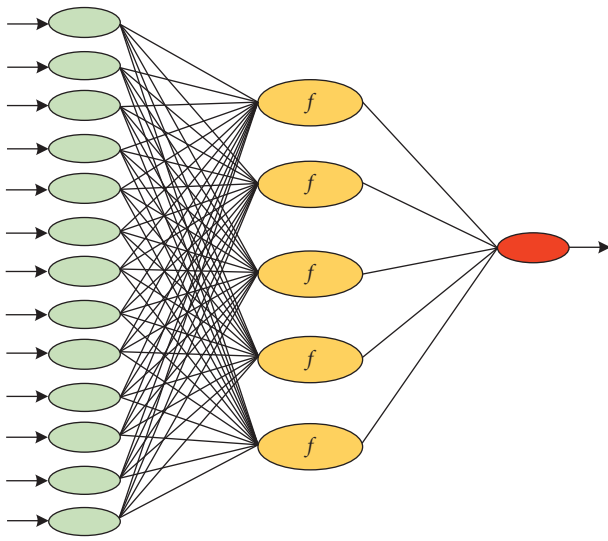


FIGURE 5: The network structure.

The study found that the input variables of the BP neural network in the previous month have a very close relationship with the input variables in the next month. The calculation results show the following rules: using the relevant factors of consumption behavior in the previous month as the input and the sales volume in the next month as the output, the model prediction values obtained from the training set and the test set have a high degree of fit, and the prediction is more accurate. Therefore, the data of November and December of each product are removed, and the output result of each data is moved up by one. That is, the input data of the previous month are used to match the output results of the next month. This also solves the problem that the unknown month cannot be obtained by the input variable. Due to insufficient inventory, some items have preorders in certain months. During data processing, the sales and preorders are aggregated, and the sum is calculated as the output result of this month.

This article uses the data from January 2017 to April 2019 as the training set and the data from May to August 2019 as the test set. After model calculation, the output result in September is the sales volume in October 2019, that is, the quantity that needs to be purchased in September. Since the purchase is made once a month, the goods purchased in September will arrive in October at the latest, which can also effectively reduce the problem of inventory backlog. The categories of predicted commodities are selected as food, cosmetics, and clothing. The normalized frequency is plotted in Figure 6.

4. Empirical Analysis

This section will perform data preprocessing on the raw data exported from the company database. Data preprocessing is the process of preprocessing dirty data such as outliers and missing values in the data set before building model features. The quality of data preprocessing directly affects the fit and robustness of the model. There are still many unreasonable parts in the current data set that need to be preprocessed before constructing features, including the processing of missing values and the smoothing of sales.

In the feature engineering of data mining, the processing methods of missing values mainly include discarding method and filling method. For features with more missing values, the discarding method is generally adopted directly, otherwise it may cause more noise, which will affect the model parameter offset; for the features with fewer missing values, the filling method is generally adopted. Generally, there are three main ways to fill missing values: one is to directly regard the missing value as an attribute feature; the other is to use a representative such as 0 or -1 that does not conflict with other attribute values; and the third is to use the mean, median, etc. to fill in; use the data above and below the missing value. The data are plotted in Figure 7.

According to the analysis of commodity information, in the current data set, there are 85.73% missing values in the seasonal attribute of commodity information, and this

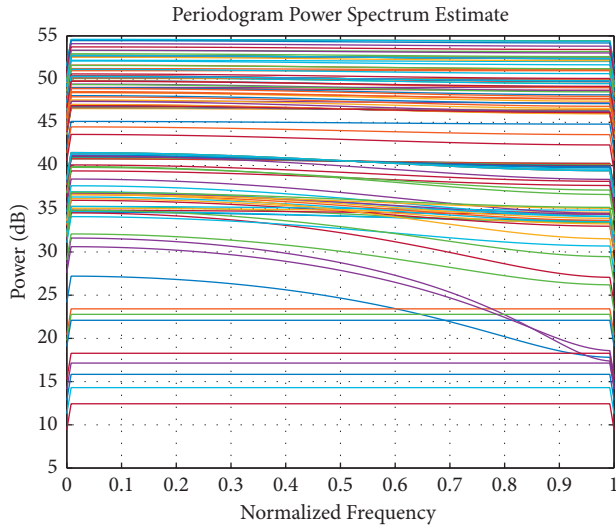


FIGURE 6: Normalized frequency.

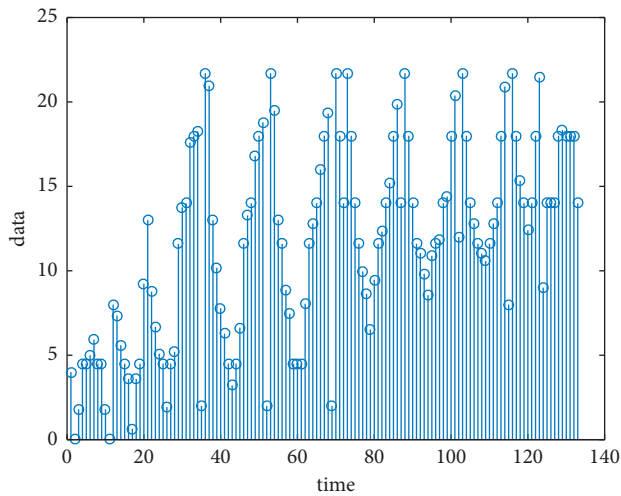


FIGURE 7: Data.

feature does not have much significance for the overall sales of commodities, so this feature is discarded. At the same time, the fifth-level, sixth-level, and seventh-level categories in the category features of commodity information have no value at all, and the method of discarding such features is directly used for processing, while the third-level category and the fourth-level category have relatively few missing values, indicating that the category division of some products has not been divided into four-level categories, and only the first few categories can be used to represent, so you can fill in these two features and other attribute values that do not conflict with -1 to handle missing values. And in the user performance data of the product and the date in the sales data of the product, there are some missing values. By comparing the upper and lower data of the missing value in the user performance data of the product, it is found that in the data with missing date, the date of the previous item, and the date of the next item are consecutive, indicating that these null data are unreasonable data generated by the

program and should delete these data directly. In the commodity sales data, the dates in the data are continuous, and the missing date data can be directly filled with the date of the previous data.

Kalman smoother calculates the current optimal amount based on the current sales volume and the previous sales volume and “error” to smooth the time series sales, eliminate the interference of “noise,” and is conducive to the learning of the model. This article uses the KalmanFilter smooth function included in Python’s pykalman library to smooth the sales data, but because the sales data are both a feature and a predicted value, this article adds a feature column to the smoothed value to avoid affecting the authenticity of the predicted value.

Data normalization is to map feature variables between 0 and 1 according to a certain scaling ratio to avoid problems caused by different dimensions. According to the user behavior characteristic variables for the weekly statistics of products, the number of sku clicks is already two orders of magnitude difference between adding to the shopping cart and adding to the collection, and the value range of the discount is between 0 and 1. If the data are not normalized, normalization processing, the difference between different dimensions, will cause the model to be more affected by the features with larger values, so it is necessary to normalize the feature variables, and process the dimensional feature variables into dimensionless feature variables.

Use the Sklearn.preprocessing.MinMaxScaler function used to normalize 6 variables such as the number of clicks, the number of add-ons, the number of favorites, the average sales volume of the brand, the average sales volume of the category, and the historical sales volume of sku and save the maximum value of each feature value and minimum value. When the model is deployed online, the same measures are also taken for the corresponding feature variables in the new data.

Using the data from 2017 to 2019 as the modeling basis, predict the subsequent data in 2020. According to the above learning steps, after thousands of times of learning, the squared error (input is the normalized sample) can be achieved. Judging from the learning speed, its speed can be compared with the commonly used BP learning algorithm. In order to verify whether the evolutionary strategy will fall into the local extreme point, 10 groups of weights were randomly selected for learning. For these 10 groups of weights, the evolutionary strategy can be successfully learned without falling into the local extreme point. Compared with the BP learning algorithm, which is easy to fall into local extreme points, it has great advantages. After learning, the network approximates all the sample points well, that is, the neural network has well realized the function mapping relationship between sales volume, price, and income. The neural network has good associative ability, so it should have good prediction ability for future sales. In order to verify this, the learned neural network is used to predict the purchase volume of cross-border e-commerce platforms after 2020. The results are shown in Figure 8. Curve in the figure is the actual purchase volume of the cross-border e-commerce platform, and the curve is the output of the

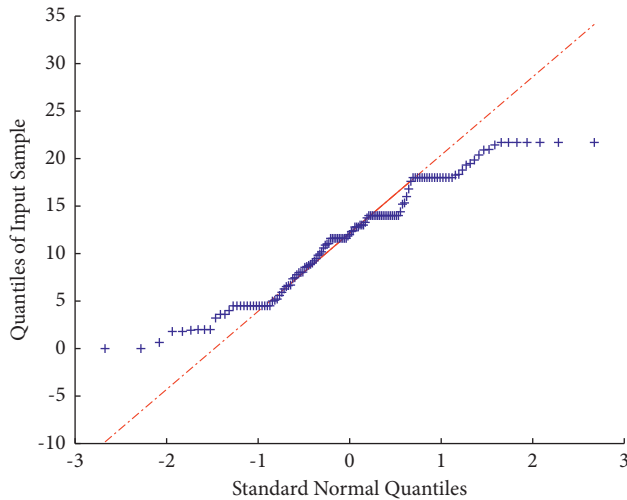


FIGURE 8: The results.

neural network. It can be seen from the figure that the neural network can well approximate the training sample points (2019) and has strong predictive power (post-2020 data). It can be seen from the figure that the average relative error of the neural network is only 1.03%, indicating that it is feasible to build a sales forecast model with a neural network, and the forecast accuracy is quite high.

Through the above comparative analysis results, it is found that the prediction error of BP neural network is small, and the prediction accuracy is relatively high. Therefore, with the help of this model, based on the data of each variable index from 2017 to 2019, the purchase volume of cross-border e-commerce platforms after 2020 is predicted and estimated. Input the standardized data of each variable index in 2018 into the already constructed cross-border e-commerce platform purchase volume prediction model (i.e. BP neural network prediction model) and finally calculate that the cross-border e-commerce platform purchase volume in the next year is approximately 1,069,940. The relative error between the sales volume predicted by the model and the actual sales volume is 1.03%. Since the prediction results of the article using the BP neural network model are based on the assumption that the country has not issued major policies for the e-commerce market and the e-commerce companies have not changed significantly, the predicted results have certain limitations, but the model the forecast results can be referenced by the power supplier company.

Based on various modeling analysis results, it can be found that the BP neural network prediction model has the advantages of high prediction accuracy and relatively accurate prediction results. Therefore, compared with the linear regression forecasting model, the BP neural network model has a broad prospect in the field of sales forecasting with many influencing factors. This model can provide an effective auxiliary means for e-commerce companies to forecast and estimate sales. E-commerce companies can use the data analysis method of BP neural network to forecast and estimate future sales, and then companies can make corresponding effective strategies with reference to the forecast results. The evaluation is shown in Figure 9.

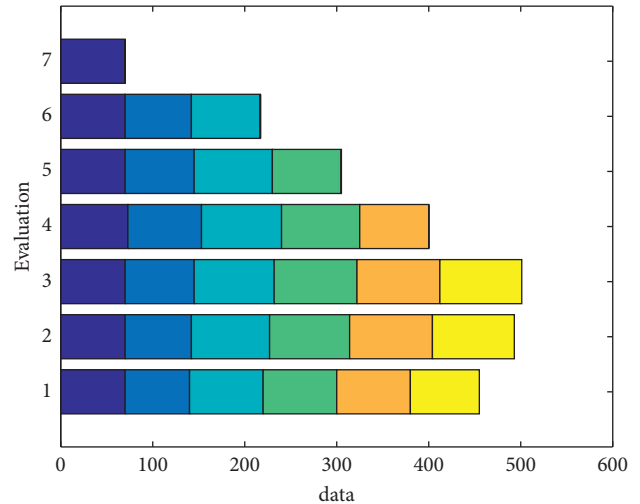


FIGURE 9: Evaluation.

5. Conclusion

This paper predicts the purchase volume of goods on an e-commerce platform. In the research, the time series characteristics of commodity sales are analyzed, and combined with commodity consumption behavior, the prediction application of BP-based neural network in cross-border e-commerce platform is proposed. Through analysis, irrelevant factors were removed, and corresponding data processing was done. Afterwards, the analysis and comparison of predictions are made with various models, and it is concluded that this method can better solve the problem of predicting the purchase volume of the platform. After verification, the average error percentage of this method is 5.9%, and the method can be extended to the practical application of enterprises.

Data Availability

The data set used in this study can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] Y. X. Xie, Y. J. Yan, X. Li, T. S. Ding, and C. Ma, "Fault diagnosis method for scintillation detector based on BP neural network," *Journal of Instrumentation*, vol. 16, no. 7, p. 403, 2021.
- [2] Y.-R. Li, T. Zhu, S.-N. Xiao et al., "Application of the collision mathematical model based on a BP neural network in railway vehicles," *Proceedings of the Institution of Mechanical Engineers - Part F: Journal of Rail and Rapid Transit*, vol. 235, no. 6, pp. 713–725, 2021.
- [3] M. Yuan and C. Li, "Research on global higher education quality based on BP neural network and analytic hierarchy process," *Journal of Computer and Communications*, vol. 9, no. 6, pp. 158–173, 2021.

- [4] C. Wang, M. Li, R. Wang, H. Yu, and S. Wang, "An image denoising method based on BP neural network optimized by improved whale optimization algorithm," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 1, pp. 1–9, 2021.
- [5] W. Wang, Y. Liu, F. Bai, and G. Xue, "Capture power prediction of the frustum of a cone shaped floating body based on BP neural network," *Journal of Marine Science and Engineering*, vol. 9, no. 6, p. 656, 2021.
- [6] J. Li, "Construction of legal incentive evaluation model based on BP neural network with multiple hidden layers," *Journal of Physics: Conference Series*, vol. 1941, no. 1, pp. 98–103, 2021.
- [7] W. Zhou, G. Sun, Z. Yang, H. Wang, L. Fang, and J. Wang, "BP neural network based reconstruction method for radiation field applications," *Nuclear Engineering and Design*, vol. 380, Article ID 111228, 2021.
- [8] A. Gupta, R. P. Narwaria, and M. Singh, "Review on deep learning handwritten digit recognition using convolutional neural network," *International Journal of Recent Technology and Engineering*, vol. 9, no. 5, pp. 245–247, 2021.
- [9] Z. Luo, F. Huang, and H. Liu, "PM2.5 concentration estimation using convolutional neural network and gradient boosting machine," *Journal of Environmental Sciences*, vol. 98, pp. 85–93, 2020.
- [10] S. Xu, "BP neural network-based detection of soil and water structure in mountainous areas and the mechanism of wearing fatigue in running sports," *Arabian Journal of Geosciences*, vol. 14, no. 11, pp. 1–9, 2021.
- [11] Z. S. Peng, H. C. Ji, W. C. Pei, B. Y. Liu, and G. Song, "Constitutive relationship of TC4 titanium alloy based on back propagating (BP) neural network (NN)," *Metalurgija*, vol. 60, no. 3–4, pp. 277–280, 2021.
- [12] H. Gao, X. Li, Z. Li et al., "The buildup factor calculations of concrete with different proportions of CRT based on a BP neural network by MCNP," *Journal of Nuclear Science and Technology*, vol. 58, no. 6, pp. 667–675, 2021.
- [13] M. A. Al Aziz, S. Mahmud, Md. A. Islam, J. Al Mahmud, and K. M. Hasib, "A comparative study of AHP and fuzzy AHP method for inconsistent data," *International Journal of Sciences: Basic and Applied Research*, vol. 54, no. 4, pp. 16–37, 2020.
- [14] Q. Liao and M. Shao, "Discussion on payment application in cross-border E-commerce platform from the perspective of blockchain," *E3S Web of Conferences*, vol. 235, Article ID 03020, 2021.
- [15] Y. Yang, L. Yang, J. Yang, C. Fan, and H. Chen, "Risk factors of consumer switching behavior for cross-border e-commerce mobile platform," *International Journal of Mobile Communications*, vol. 18, no. 1, p. 1, 2020.
- [16] S. Ding, Y. Hu, and Z. Xiong, "Research on online discourse of cross-border E-commerce platform on the basis of Co-operative principle," *Studies in Literature and Language*, vol. 21, no. 2, pp. 37–44, 2020.
- [17] J. Mou, J. Cohen, Y. Dou, and B. Zhang, "International buyers' repurchase intentions in a Chinese cross-border e-commerce platform," *Internet Research*, vol. 30, no. 2, pp. 403–437, 2019.
- [18] J. Liu, R. Zhang, G. Han, N. Sun, and S. Kwong, "Video action recognition with visual privacy protection based on compressed sensing," *Journal of Systems Architecture*, vol. 113, Article ID 101882, 2020.
- [19] X. Wu and Q. Ji, "TBRNet: Two-stream BiLSTM residual network for video action recognition," *Algorithms*, vol. 13, no. 7, pp. 1–9, 2020.
- [20] Lv Qin, "Supply chain decision-making of cross-border E-commerce platforms," *Advances in Industrial Engineering and Management*, vol. 7, no. 1, pp. 1–8, 2018.
- [21] Y. Chung, "Artificial intelligence and the virtual multi-door ODR platform for small value cross-border e-commerce disputes," *Journal of Arbitration Studies*, vol. 29, no. 3, pp. 99–119, 2019.
- [22] J.-W. Oh, B. Huh, and M.-R. Kim, "Effect of learning contracts in clinical pediatric nursing education on students' outcomes: A research article," *Nurse Education Today*, vol. 83, Article ID 104191, 2019.
- [23] F. Aulkemeier, M.-E. Iacob, and J. van Hillegersberg, "An architectural perspective on service adoption: A platform design and the case of pluggable cross-border trade compliance in e-commerce," *Journal of Organizational Computing & Electronic Commerce*, vol. 27, no. 4, pp. 325–341, 2017.
- [24] W. Wei, X. Xia, M. Wozniak, X. Fan, R. Damaševičius, and Y. Li, "Multi-sink distributed power control algorithm for cyber-physical-systems in coal mine tunnels," *Computer Networks*, vol. 161, pp. 210–219, 2019.
- [25] Z. Huangfu, H. Hu, N. Xie, Y.-Q. Zhu, H. Chen, and Y. Wang, "The heterogeneous influence of economic growth on environmental pollution: Evidence from municipal data of China," *Petroleum Science*, vol. 17, no. 4, pp. 1180–1193, 2020.

Retraction

Retracted: Intelligent Recognition Model of Business English Translation Based on Improved GLR Algorithm

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

References

- [1] L. Deng, X. Hu, and F. Liu, "Intelligent Recognition Model of Business English Translation Based on Improved GLR Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4105942, 8 pages, 2022.

Research Article

Intelligent Recognition Model of Business English Translation Based on Improved GLR Algorithm

Laiying Deng,¹ Xinde Hu,² and Fen Liu ¹

¹Guangzhou City Polytechnic, Guangzhou, Guangdong 510000, China

²Beijing Qingneng Internet Technology Co., Ltd, Guangzhou Branch, Guangzhou, Guangdong 510000, China

Correspondence should be addressed to Fen Liu; lfcathie@gcp.edu.cn

Received 25 February 2022; Accepted 31 March 2022; Published 15 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Laiying Deng 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 problem of low accuracy of traditional algorithm model, an intelligent recognition model of business English translation based on an improved GLR algorithm is proposed. Through this algorithm, the automatic sentence recognition technology is established, and according to the characteristics of business English, the improved GLR algorithm is used for collection, sorting, and analysis, so as to realize the intelligent recognition of business English. The results show that based on the improved GLR algorithm, the recognition accuracy is high, and the comprehensive score is 92.5 points, which overcomes the disadvantages of the GLR algorithm, and the operation speed and processing are improved. Based on the improved GLR algorithm, the intelligent translation of business English is realized, which is accurate and fast, and greatly promotes the learning and development of business English.

1. Introduction

After joining the WTO, China's international business exchanges have become more frequent, the foreign trade industry has developed rapidly, and the demand for business English talents is increasing day by day. In this context, more and more colleges and universities begin to set up business English majors [1]. Business English translation course is an important skill course of this major, focusing on translation teaching. International business communication is mostly completed through language communication, and one of the common languages is English. The main communication language between Chinese businessmen and foreign countries is English. Therefore, the importance of business English translation teaching is self-evident. Recently, as the science and technology develop, the artificial intelligence is more and more widely applied, and machine translation is constantly updated and upgraded, which poses a challenge to the traditional human translation. There are some new problems in the teaching of business English translation in Colleges and universities, mainly manifested in the lagging consciousness of teachers, the outdated content of teaching materials and the single teaching methods [2–4].

As far as teachers' awareness is concerned, under the current situation that the functions of translation software are constantly upgraded and updated, many teachers do not understand the advantages and disadvantages of machine translation timely and fully. They can neither make use of its advantages in teaching nor face the challenges it brings, so as to adjust the teaching content accordingly [5]. As far as the content of teaching materials is concerned, although the quantity and quality of business English translation teaching materials on the market have improved in recent ten years, due to the limited factors such as long publishing cycle, limited energy of teachers, difficulty in collecting high-quality translation materials, rapid development of business society, the content of teaching materials cannot be updated in time, and the listed cases are relatively outdated. In terms of teaching methods, at present, in business English translation classes in colleges and universities, many teachers still use traditional teaching methods, focusing on theory, culture teaching, and translation practice. Sometimes, the explanation of translation theory is too much and too detailed. When arranging translation practice exercises, some translation examples are outdated, and the form and type of translation tasks tend to be single. Teachers' guidance

of students' translation is not enough, so the effect of translation practice is not satisfactory, which affects the realization of the goal of talent training. It can be said that teaching methods have not been adjusted and improved according to the development, changes and trends of the times, society, and translation. In a word, the teaching of business English translation course cannot advance with the times, which is disconnected from the needs of society and industry. Therefore, there is an urgent need for reform in all aspects [6–8].

Machine translation, also known as automatic translation or computer translation, begins to develop in the 1940s. After years of development, it has gradually transformed from a very optimistic automatic high-quality machine translation to a more realistic and feasible computer-assisted translation. We can be seen the ideal of realizing unmanned translation was born as early as the 1940s, and CAT evolved from the original machine translation. During the development of computer-assisted translation, major business giants have also launched automatic translation services, such as Google translate and Baidu translate. In 2017, the artificial intelligence translation box appeared, and the speed and accuracy of the product's bilingual translation of everyday words are amazing. Instant artificial intelligence translation becomes a hot topic of discussion.

Recently, with the rapid development of intelligence, the use of computer technology to translate English has become the mainstream of the times. Now, there is a large market demand. Companies, universities, and professionals are like to use intelligent translation technology to translate business English, which brings great convenience and reveals that some machine translation levels have been unable to meet the current translation needs [9]. The market urgently needs a machine translation technology with high performance and high translation accuracy [10, 11]. The core idea of computer-aided translation is the translations are usually regarded as auxiliary references, so the users could obtain the advantages of translation and make manual selection. On the other hand, the use of corpus can classify and sort out the vocabulary in various industries, so that the quality of translation can be improved and closer to the actual needs of users. The rational use of professional vocabulary corpus with high translation frequency can not only greatly reduce repetition and translation workload but also greatly improve the accuracy of translation [12].

With the help of big data, the translation results of artificial intelligence translation software can be obtained instantaneously, which is conducive to learners' more efficient learning and consolidate their own learning effect [13, 14]. In other words, if they do not review after learning, the longer it takes, the less memory, or even forgetting. Learning needs to be reviewed constantly. By the artificial intelligence translation, we can learn faster and review better. No matter when and where we don't know, we can check if we are uncertain and review if we think of it. With the timeliness of the Internet, we can carry out more meaningful learning, reduce mechanical learning, and improve learning efficiency.

Students can use online cloud translation and online electronic dictionary to assist learning and make learning

more efficient [15, 16]. Many students like to look at the word list and hold the word book to memorize words, but such learning is fragmented learning. They can only see the trees but not the forest. If they remember the words, they may not be able to use them. Learning is easy to be frustrated. However, the use of intelligent translation-assisted learning can avoid the situation of learning words without understanding the application and can get through the vocabulary more quickly and better, because intelligent translation presents learners with complete sentences rather than fragmented language. We can even make use of the advantage of intelligent translation based on big data to select corpus closer to our lives and our favorite fields for learning, so as to make learning more interesting and grounded. This is especially suitable for students with a slightly poor foundation. They can use translation software to make up for the personalized needs that cannot be met in the big classroom, improve their interest in learning, make up for their shortcomings, and make learning happen continuously.

Under the influence of the Internet and big data, the language learning model needs to be changed [17]. Foreign language learners can no longer be limited to the learning of words, phrases, and simple sentence patterns but should develop more comprehensive abilities. Compared with the past, using the mobile app to assist learning, improving listening, or writing ability by visiting web pages, and using the fragmented time to read English novels to improve reading ability, intelligent translation is more comprehensive and effective in assisting learning [18]. Students can not only master the grammar and grammar but also improve their ability to translate and use the words intelligently.

Zhou [19] analyzed the detailed features of machine translation in Business English, adopted the NT clause as the unit in English translation, and then achieved the purpose of business English translation through the PTA model, which reflected the intelligence of small languages. Lu [20] improves the traditional intelligent recognition model of business English translation and adopts the language intelligent translation model. Firstly, the sinusoidal similarity algorithm is used to obtain the measurement values of different semantics, so as to obtain the accuracy of business English translation and ensure the accuracy of business English. Huang [21] used the analytic hierarchy process to analyze the business English translation and made a sentence comparative analysis of the English Thesaurus and the cut words, so as to obtain the possible grammatical composition of business English, reduce the error rate of business English translation, and improve the accuracy and accuracy of the translation.

In general, this paper implements a sentence tagging system suitable for small and medium-sized syntactic tagging corpus method analyzers. In the implementation process, in order to improve the sensitivity of lexical annotation information to differences, the number of rules shall be reduced as far as possible, and the text shall be marked at the first level, but the verbs shall be marked. At the same time, some auxiliary words and punctuation in sentences are labeled in detail, so that some ambiguities can be

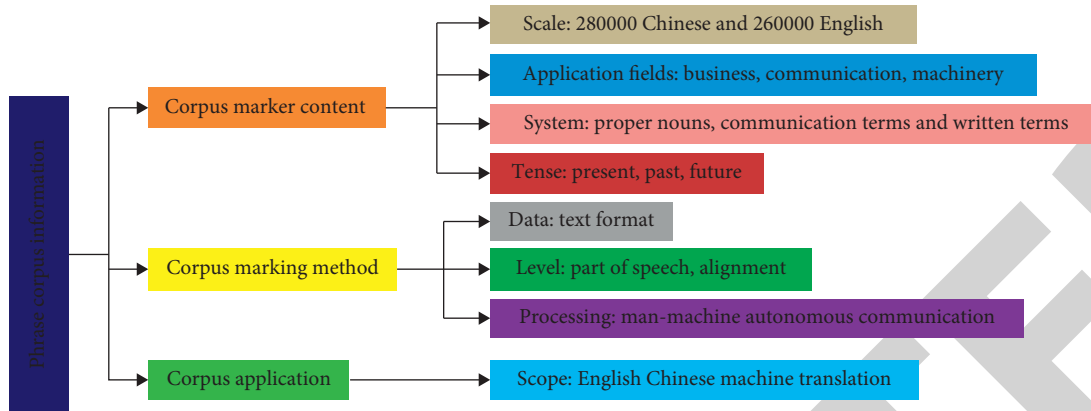


FIGURE 1: Phrase corpus information flow.

eliminated through detailed part of speech labeling information in the ruling stage [22]. Structural ambiguity is a difficulty in the field of English translation, which needs to be solved by part of speech recognition algorithm. In order to use English in daily communication practice, we should strengthen English pronunciation correction and learning. In the past, it was completely taught by teachers, which not only required a lot of human resources but also difficult to realize. Therefore, it is particularly important to develop an intelligent recognition model for English translation. The model can monitor students' English pronunciation and put forward correction suggestions. There are some problems in the traditional intelligent recognition model and setting of the English translation, which cannot accurately recognize English translation and cannot correct students' English pronunciation, but also mislead students' pronunciation and affect English learning.

Based on the above problems, this paper establishes a language database, improves and analyzes the GLR algorithm, establishes an improved GLR algorithm model, and achieves the function of intelligent recognition. According to the syntactic function of parsing the linear table, this paper corrects the English Chinese structural ambiguity in the part of speech recognition results, finally obtains the recognized content, and determines the actual position range of phrases in translation, in order to alleviate the disadvantages of structural ambiguity in the current field of the English translation to a certain extent and to improve the efficiency of phrase recognition. Finally, an intelligent recognition model of business English translation based on an improved GLR algorithm is established to automatically recognize English translation. In order to verify the effectiveness of the model, simulation experiments are carried out in the simulation environment.

2. English Intelligent Recognition Algorithm

2.1. Creating Phrase Corpus. In Chinese, words are divided into notional words and functional words. Notional words refer to those components that express real meaning, can be used as phrases or sentences, and can form sentences alone. Function words generally do not express real meaning and

are not used as components of phrases or sentences (except adverbs). Their basic purpose is to express grammatical relationships. In order to obtain as many analysis results as possible, thread technology is adopted to display all the possible analysis results of one sentence for the probability model to choose the best one [23]. Figure 1 shows the information flow of the phrase corpus.

The English translation intelligent recognition model designed in this paper has more than 500000 words in the phrase corpus, which can meet the needs of practical use. As shown in Figure 1, the phrase corpus is highly targeted. Through the English Chinese machine translation phrase corpus, this paper comprehensively distinguishes the tenses of different phrase corpora and labels the English Chinese phrase corpus. The corpus marking method includes three parts: level, data, and processing. The data type is text format. The level uses alignment and part of speech.

2.2. Phrase Corpus Part of Speech Recognition. Noun phrase is a very important phrase type, which accounts for a large proportion in Chinese sentences. Recognizing short nouns is a very important subtask in the field of natural language processing. It is directly related to the correctness of text analysis and text processing. As a preprocessing method of natural language processing, its analysis results can simplify the sentence structure and solve the ambiguity problem, so as to reduce the difficulty and complexity of syntactic analysis and provide the basis for further phrase division and syntactic analysis. At the same time, it is of great significance to the application fields of machine translation, information retrieval, information extraction, and corpus research [24].

GLR algorithm, also known as the Futian algorithm, is an extended LR algorithm. The algorithm introduces graph structure stack, subtree sharing, and local ambiguity reduction. In order to make the space complexity as small as possible in the process of constructing the item set family of the algorithm, the merging strategy of item set is adopted; In order to obtain as many analysis results as possible, thread technology is adopted to display all the possible analysis results of one sentence for the probability model to choose the best one.

The operation of each step of the classical GLR algorithm uses a variety of shift instructions and simplified operations.

The beginning and terminal of each operation during the period are displayed using the special envoy's standard. In the process of phrase translation, when the GLR algorithm does not detect grammatical ambiguity, it will restart the operation of deduplication and calibration. If grammatical ambiguity is detected, it is necessary to use the geometric structure linear table of syntactic analysis, which provides the optimal content according to the local optimization principle, transports it to different recognition channels for symbol recognition, and selects the optimal result according to the recognition result [25].

Generally, the traditional GLR algorithm has a low degree of character recognition, so it cannot effectively translate business English and apply more accurate English translation. Based on the improved GLR algorithm, this paper establishes the automatic recognition ability, which greatly improves the accuracy of sentence recognition as shown as follows:

$$G_E = (V_N, V_T, S, \alpha). \quad (1)$$

If P represents α any action in and exists in V_N , such as language translation, grammar checking, speech recognition, and other actions. Through derivation, it can be concluded that

$$P \longrightarrow \{\theta, c, x, \delta\}. \quad (2)$$

When doing the GLR based parsing algorithm, due to the splitting of some concentric sets in the process of constructing the state transition graph, the number of result states increases sharply, resulting in a sharp increase in the complexity of program space and a large amount of memory space. When there are many rules, the execution of the program will overflow the memory [26].

3. English Translation Intelligent Recognition Model

The intelligent recognition model is planned according to data collection, processing, and output. Figure 2 shows the overall architecture of intelligent recognition of business English translation.

3.1. Model Design Process. The functions are planned required by the English translation intelligent recognition model, and the overall model is designed. Figure 3 shows the model design process. This model can realize data collection, output, and processing. The data acquisition device is used to collect the voice signal, and then, the audio input device is used to input the English signal into the processing system to process the data signal. The processed results are output in the corresponding client and displayed. The user can view the automatic recognition results of English translation through the display or client.

3.2. English Signal Processing. After the model design, we should carry out a detailed design, collect, and process English signals in a planned way. However, because English speech signals have interference factors, such as

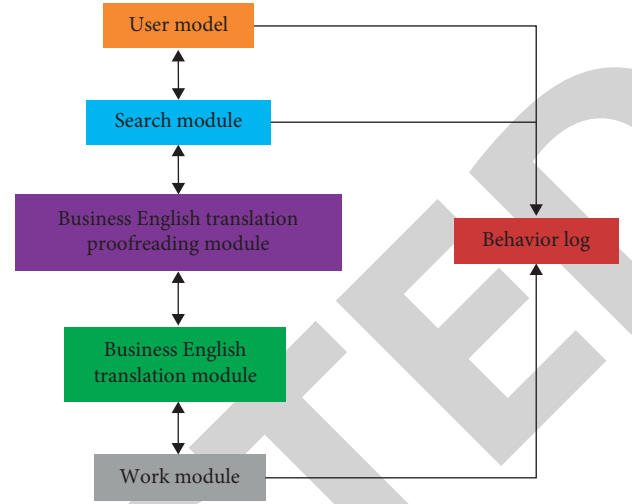


FIGURE 2: General framework of intelligent recognition of business English translation.

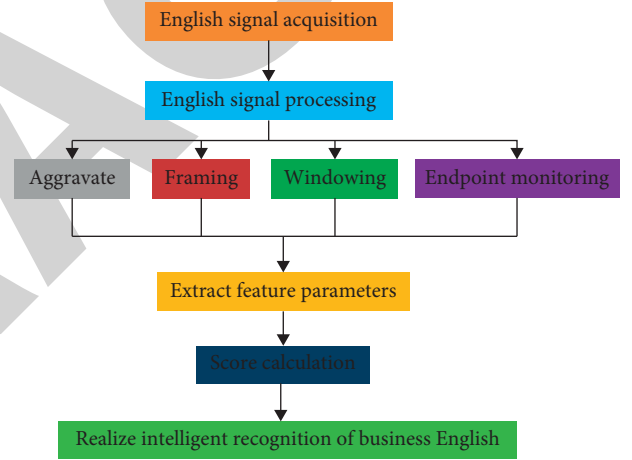


FIGURE 3: Model design process.

nonstandard pronunciation of the human body and imperfect equipment, the collected speech signal information is not 100% accurate. In order to improve the accuracy, the collected voice signals should be processed, including framing, weighting, endpoint monitoring, and windowing.

The digital filter is used to realize the speech signal weighting processing and improve the stress detection system. The specific method is to use the symbols $F1$ and $F2$ to represent the first and second formants of vowel spectrum characteristics, respectively, which are related to the front and back of tongue position and high and low. Based on this, after using a neural network to distinguish nonstress and stress, we can use the classifier system to output position confidence, so as to measure the vowel intonation, and then select the best speech signal. In this paper, the high-frequency signal lifting pre-emphasis digital filter with a signal-to-noise ratio of 6 dB is used to process the frequency band of speech signal higher than 800 Hz. Weighting coefficient α between 0.9 and 1.0, the calculation formula of weighting processing signal $y(n)$ is

$$\begin{aligned} y(n) &= T[x(n)] \\ &= ax(n) + b \text{ (} a \text{ and } b \text{ are constants and not 0),} \end{aligned} \quad (3)$$

where $x(n)$ is the input voice signal. Among them, a can be taken according to the high-frequency signal, and b is the frequency band value, which can be obtained according to the experimental results. After signal processing, the voice signal is processed in the second step, and the half-frame overlapping method is used to realize the framing operation. In order to accurately analyze the speech signal, the speech signal is divided into t frames, and the calculation formula is

$$z(n) = \frac{1}{t} y(n), \quad (4)$$

where $z(n)$ is the speech signal after dividing the t frame. After the framing operation processing is realized, the voice signal is windowed. In order to clearly show the voice effect, the rectangular window $w(n)$ is selected, and the calculation formula is

$$w(n) = \omega(n) \times z(n). \quad (5)$$

$\omega(n)$ is the windowed speech signal, which has a fast response and improves the translation accuracy. The end-point monitoring of processed speech signal is realized by using the double threshold comparison method, so as to quickly find the starting point and ending point for data processing and storage. This method can reduce the interference items and improve the monitoring accuracy.

3.3. Extracting Characteristic Parameters. In order to train the later feature extraction and parameter estimation, the corpus must be preprocessed first. Each word is extracted in the text and its context information to form a word item, and the frequency of the same word item is recorded. All word items form a thesaurus containing all the characteristic attributes of each word and its context.

In the test model, the input text is preprocessed to obtain the feature information of each word and its context, and then, the feature vector satisfied by the word is calculated; finally, the feature vector and feature parameters are used to label the right part of the feature, and the maximum labeling result is the final labeling result. Through the improvement of a finite length discrete speech signal, the following formula is obtained:

$$X[K] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}nk}, \quad k = 0, 1, 2, \dots, N, \quad (6)$$

where $x[n]$ is the discrete speech sequence obtained by sampling, and $X[K]$ is the k -point reset sequence.

Discrete speech sequences are transformed into Mel frequency scales.

$$\text{Mel}(f) = 2579 \lg\left(1 + \frac{f}{700}\right), \quad (7)$$

where $\text{Mel}(f)$ is the Mel frequency and f is the actual frequency.

Discrete cosine transform DTC on the filtered output is carried out to obtain the feature parameter extraction result P of speech signal $w(n)$. The calculation formula is

$$P = Z_{n=1}^N F(l)w(n)\cos(\pi n(M + 0.5)). \quad (8)$$

When an improved GLR algorithm is used to analyze sentences, it will first reduce the complexity of sentences and then adopt automatic recognition technology to translate business English through the function of its own algorithm. Moreover, under the intelligent recognition technology, using the above method, taking English speech intelligent recognition as the demand, the relevant speech signal parameters are extracted from the aspects of rhythm, speed, intonation, and intonation.

4. Test and Model Verification

4.1. Verification Scheme. In order to verify the actual English Chinese translation effect of the improved GLR algorithm, based on the above methods, this paper preliminarily realizes a prototype system and preliminarily carries out some experiments for the word alignment method based on dictionary and semantic similarity, the word alignment method based on statistics, and the method based on the integration of statistics and dictionary to test the effectiveness of different word alignment methods. The influence of noun phrase alignment based on word alignment position information is proposed in this paper. Afterwards, the dictionary-based method combines the advantages of the statistical method and the dictionary-based method, which not only makes up for the lack of accuracy in the statistical method, increases the number of correct alignment, and ensures the accuracy of nonempty alignment but also overcomes the problem of limited dictionary coverage in the dictionary-based method, and further improves the translation accuracy of business English.

4.2. Experimental Results. Through the intelligent recognition technology of business English, the results are shown in Figure 4. It can be seen that the improved GLR algorithm for part of speech recognition is the best in the same category, followed by the GLR algorithm, and the worst is the statistical algorithm. From Figure 5, we can find that the improved GLR algorithm can obtain the high score, which is 92.5 points, and the statistical algorithm is 76.0 points, which is the lowest. Figure 6 shows the frequency intensity analysis under different algorithms. It can be seen that the frequency intensity of the improved GLR algorithm is relatively stable and close, while the intensity of the statistical algorithm under different frequencies is poor, which shows that the translation accuracy of the improved GLR algorithm is the highest. In general, machine translation based on the statistical algorithm and dynamic memory algorithm did not translate the word "Price Bureau," but machine translation based on improved GLR algorithm correctly translated it. When translating the "explanation," only the machine translation based on the improved GLR algorithm is closest to the human translation. It can be seen that the machine

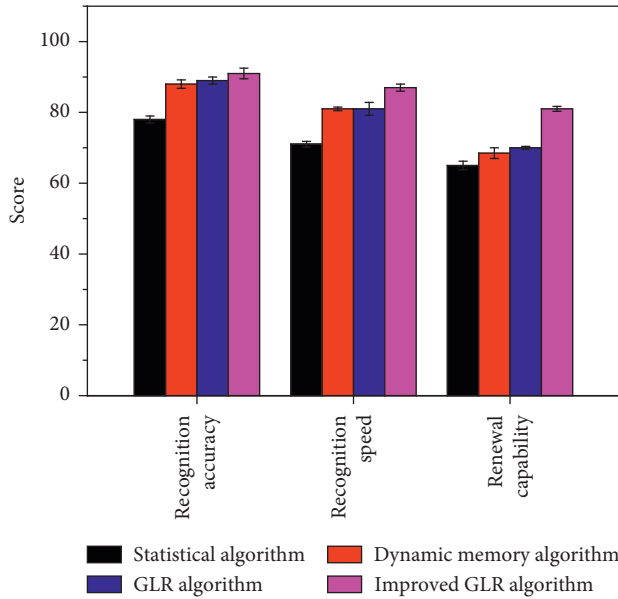


FIGURE 4: Evaluation results of four English Chinese translation algorithms.

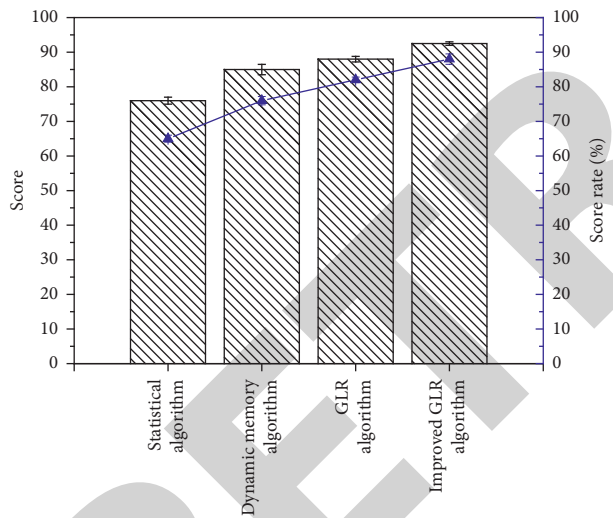


FIGURE 5: Comparison of comprehensive test scores of four English Chinese translation algorithms.

translation based on the improved GLR algorithm designed in this paper is more accurate than the statistical algorithm and the dynamic memory algorithm, reaching the same level as the human translation. Chinese Grammar presents integrity, fuzziness, and high flexibility; English grammar is highly systematic. Due to the limited vocabulary of learners, their understanding of language knowledge is not comprehensive. The mastery of grammar rules is ambiguous and limited by the Chinese way of thinking. Their foreign language learning generally has a thinking transformation process from English to Chinese and then to English. This kind of error is almost difficult for learners to detect without teachers' guidance, so teachers should point out and correct this kind of error. It shows that the improved GLR algorithm

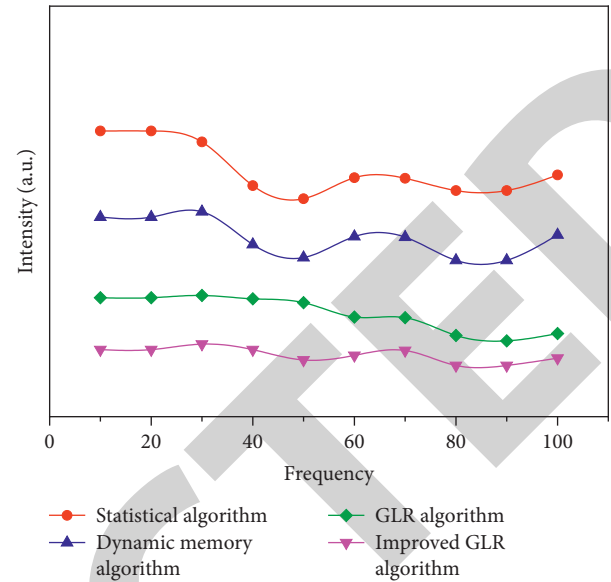


FIGURE 6: Frequency intensity analysis under different algorithms.

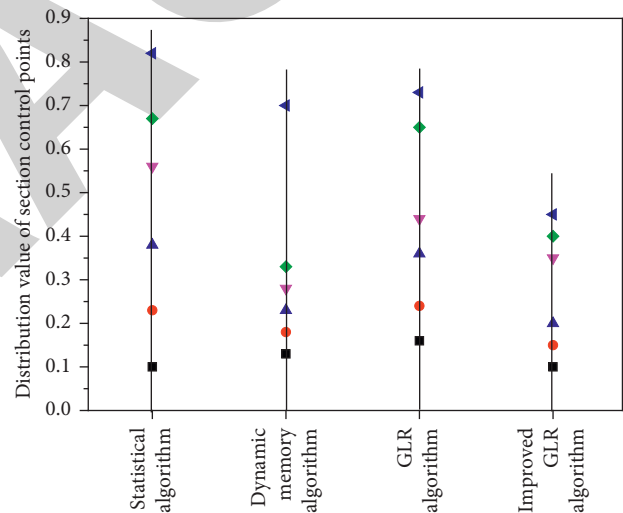


FIGURE 7: The distribution of control points of the system.

is efficient and feasible in machine translation. Combined with Figures 4 to 6, the improved GLR algorithm has a higher score and better intelligent pattern recognition.

Figure 7 shows the distribution of section control points identified by the system. It can be seen that the syntax and phrase recognition system obtained by the statistical algorithm has loose node distribution, while the syntax and phrase obtained by the dynamic memory algorithm and GLR algorithm have a compact node distribution, indicating that the intelligent recognition technology of the representation model is high and the translation accuracy is good. The compact distribution of node control points obtained by the improved GLR algorithm shows that the system is easy to operate and has high precision, which is helpful for business people to use, and can promote the rapid development of business English translation.

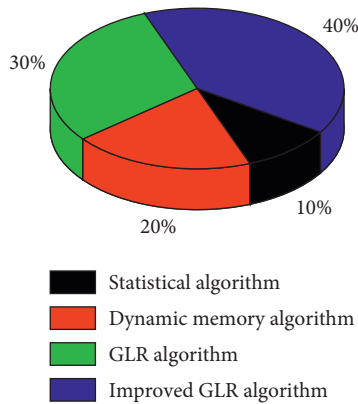


FIGURE 8: Proportion of different algorithms in business English translation.

Each algorithm has its unique advantages. In this paper, the applicable places of various algorithms are discussed in detail and then applied to the teaching management of business English. Its advantages and disadvantages are compared, so as to better highlight the advantages of an improved GLR algorithm. Figure 8 shows the proportion of different algorithms in Business English translation. It suggests that the improved GLR algorithm is the most widely used, accounting for 40%, followed by the GLR algorithm, and the statistical algorithm is the lowest, accounting for only 10%. This shows that the improved GLR algorithm is the most commonly used in Business English translation, which is consistent with the above analysis results. The application potential of the improved GLR algorithm in Business English intelligent translation is unlimited. For business English in high-frequency compound noun phrases, many statistical indicators can be obtained effectively. For low-frequency compound noun phrases, less language phenomena can be observed in the corpus, and various statistical indicators do not perform well. Therefore, based on the improved GLR algorithm, high-frequency compound noun phrases are used to help find low-frequency compound noun phrases, so as to achieve the results of intelligent automatic translation of business English.

5. Conclusion

Aiming at the structural problems existing in English translation, an improved GLR algorithm is proposed to solve the data point coincidence problem in the traditional GLR algorithm. The phrase structure in the improved algorithm is designed by using the phrase center point to analyze the syntactic function of the linear table, correct the ambiguity of part of speech recognition structure and English Chinese structure, solve the problem of low accuracy of recognition results in traditional calculation methods, and provide a reasonable method for phrase recognition. The experimental results show that compared with other algorithms, the machine translation based on the improved GLR algorithm has the characteristics of simple and fast calculation, low difficulty, and stronger practicability. In general, the improved

GLR algorithm meets the needs of English machine translation and is suitable for English machine translation.

Data Availability

No data were used.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The authors thank Guangzhou Educational and Scientific Planning Project in 2021, for the Integration of “double innovation” into the Higher Vocational Business English Teaching System (no. 202113570).

References

- [1] J. Cao, C. Song, S. Peng, F. Xiao, and S. Song, “Improved traffic sign detection and recognition algorithm for intelligent vehicles,” *Sensors*, vol. 19, no. 18, p. 4021, 2019.
- [2] Q. Ding and Z. Ding, “Machine learning model for feature recognition of sports competition based on improved TLD algorithm,” *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 1, pp. 1–12, 2020.
- [3] L. Lin, J. Liu, X. Zhang, and X. Liang, “Automatic translation of spoken English based on improved machine learning algorithm,” *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2385–2395, 2021.
- [4] A. Borsotti, L. Breveglieri, S. Crespi Reghizzi, and A. Morzenti, “Fast GLR parsers for extended BNF grammars and transition networks,” *Journal of Computer Languages*, vol. 64, no. 8, Article ID 101035, 2021.
- [5] M. Ghasemi, A. Sheikhi, and M. M. Pishrow, “Multiple radar subbands fusion technique based on generalized likelihood ratio test,” *Physical Communication*, vol. 46, no. 12, Article ID 101331, 2021.
- [6] S. I. Okafor, A. H. Amadi, and M. A. Abegunde, “The choke as a brainbox for smart wellhead control,” *European Journal of Engineering and Technology Research*, vol. 6, no. 1, pp. 114–118, 2021.
- [7] T. Li, Y.-H. Chan, and D. P. K. Lun, “Improved multiple-image-based reflection removal algorithm using deep neural networks,” *IEEE Transactions on Image Processing*, vol. 30, pp. 68–79, 2021.
- [8] L. I. Null, N. Tang, and N. Yang, “An improved branching algorithm for the proper interval edge deletion problem,” *Frontiers of Computer Science in China*, vol. 16, no. 2, Article ID 162401, 2022.
- [9] L. Zhang, L. Ding, S. Ullah et al., “An improved medial axis path generation algorithm for selective laser melting,” *Rapid Prototyping Journal*, vol. 26, no. 10, pp. 1751–1759, 2021.
- [10] M. Fazlollahpoor, M. Derakhtian, and S. Khorshidi, “Passive MIMO radar detection in the presence of clutter or multi-path without reference channel,” *IET Radar, Sonar & Navigation*, vol. 15, no. 2, pp. 154–166, 2021.
- [11] S. B. Chaabane, A. Belazi, S. Kharbech, A. Bouallegue, and L. Clavier, “Improved salp swarm optimization algorithm: application in feature weighting for blind modulation identification,” *Electronics*, vol. 10, no. 16, p. 2002, 2021.

Research Article

A Novel Encoder-Decoder Model for Multivariate Time Series Forecasting

Huihui Zhang ^{1,2}, Shicheng Li ³, Yu Chen ³, Jiangyan Dai ² and Yugen Yi ³

¹School of Computer Science and Technology, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China

²School of Computer Engineering, Weifang University, Weifang, China

³School of Software, Jiangxi Normal University, Nanchang, China

Correspondence should be addressed to Huihui Zhang; huihui@wfu.edu.cn and Yugen Yi; yiyg510@jxnu.edu.cn

Received 27 February 2022; Revised 26 March 2022; Accepted 28 March 2022; Published 14 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Huihui 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.

The time series is a kind of complex structure data, which contains some special characteristics such as high dimension, dynamic, and high noise. Moreover, multivariate time series (MTS) has become a crucial study in data mining. The MTS utilizes the historical data to forecast its variation trend and has turned into one of the hotspots. In the era of rapid information development and big data, accurate prediction of MTS has attracted much attention. In this paper, a novel deep learning architecture based on the encoder-decoder framework is proposed for MTS forecasting. In this architecture, firstly, the gated recurrent unit (GRU) is taken as the main unit structure of both the procedures in encoding and decoding to extract the useful successive feature information. Then, different from the existing models, the attention mechanism (AM) is introduced to exploit the importance of different historical data for reconstruction at the decoding stage. Meanwhile, feature reuse is realized by skip connections based on the residual network for alleviating the influence of previous features on data reconstruction. Finally, in order to enhance the performance and the discriminative ability of the new MTS, the convolutional structure and fully connected module are established. Furthermore, to better validate the effectiveness of MTS forecasting, extensive experiments are executed on two different types of MTS such as stock data and shared bicycle data, respectively. The experimental results adequately demonstrate the effectiveness and the feasibility of the proposed method.

1. Introduction

Time series is the sequence of arranged numbers according to the occurrence time, which is also called dynamic series. The time span can be years, quarters, months, hours, or other factors [1]. In recent years, time series are widely applied in various fields, such as economics, medicine, transportation, and environmental science, which has been attracted much attention [2]. According to the number of observed variables, time series data can be divided into univariate time series data and multivariate time series data [2]. Therefore, how to mine useful information from these time series data becomes a very important task in data mining, machine learning, artificial intelligence, and other fields [3]. As a key and crucial branch of time series data analysis, time series prediction aims to accurately predict or estimate the future

events by exploring the past and current data of the single variable or several correlated variables [4]. The former is called univariate time series forecasting; the latter is called multivariate time series forecasting. For example, economists utilized the historical data of stock prices to forecast stock prices or trends [5], medical scientists made use of the biological time data to predict diseases [6], transportation departments explored the historical data of traffic flow to predict congestion [7], and environmentalists employed atmospheric timing data to estimate environmental climate changes [8], etc. Nevertheless, time series data not only contains abundant information but also appears to some complex characteristics such as high dimension, nonlinear, fluctuation, and spatiotemporal dependence, which make accurate time series data prediction become a challenging study hotspot [9].

In the past few decades, time series data prediction has been widely concerned and many methods have been proposed [10]. For instance, traditional statistics-based methods focused on relevant domain knowledge, while learning-based methods are introduced to learn temporal dynamics in a pure data-driven strategy. As a popular learning-based method, deep learning can learn the deep latent features from the input data comprehensively and has become a cutting-edge approach [11].

The traditional statistics-based methods include autoregressive (AR) [12], autoregressive moving average (ARMA) [13], autoregressive integrated moving average, and exponential smoothing models (ARIMA) [14]. Although the above methods can utilize statistical inference to describe and evaluate the relationship between variables, they assumed that the input data has a linear relationship between model structure and the constant variance [15]. Therefore, there are some limitations to dealing with complex time series data containing nonlinear and nonstationary structures, so they cannot effectively obtain accurate predictions.

In order to solve the shortcomings mentioned above, many learning-based methods including support vector machine (SVM) [16], genetic algorithm (GA) [17], AdaBoost [18], and artificial neural network (ANN) [19], which can simulate the complex structures of time series data, have been widely applied to time series prediction task. For example, Dong et al. [16] discussed utilizing SVM for predicting building energy consumption in tropical regions, and they considered that it was superior to other neural networks from the views of performance and parameter selection. Yadav et al. [17] proposed a neuron model based on polynomial structure and used the Internet traffic and financial time series data to conduct forecast experiments, which showed that the neural network (NN) model not only achieved better performance but also greatly reduced the computational complexity and running time comparing with the existing multilayer neural networks. However, building an effective learning-based model needs a large amount of professional data, and the training process requires a high level of computer hardware equipment. Therefore, the application of traditional machine learning models is largely limited.

In recent years, with the improvement of data acquisition and computing power, a novel learning-based method called deep learning has attracted much attention. Deep learning [20] can obtain a higher-level representation of the original input via designing simple and nonlinear modules, which was conducive to learning the feature representation. Convolutional neural network (CNN) [21], recurrent neural network (RNN) [22], and variant models have been successfully applied to time series prediction. Zhang et al. [23] proposed a deep spatiotemporal residual network model to predict the flow of people throughout the city. Jagannatha and Yu [24] developed a bi-directional recurrent neural network (BRNN) for medical events detection in electronic medical records. Nevertheless, RNN and BRNN are easy to suffer from the gradient vanishing and gradient exploding problems. To overcome the drawbacks, the long short-term memory network (LSTM) [25] and the gated recurrent unit (GRU) [26] were developed. Since both LSTM and GRU can keep the historical information for a

longer time step, they are widely used in time series data analysis, prediction, and classification tasks. Compared with LSTM, the GRU has a simpler structure and fewer parameters, which can reduce the overfitting risk. For example, Shu et al. [27] presented a new neural network model based on improved GRU to predict short-term traffic flow.

As an unsupervised method, Autoencoder (AE) is also widely applied to feature representation learning [28]. In order to extract better features, the RNN is frequently combined with AE. Xu and Yoneda [29] first used a stacked autoencoder (SAE) to encode the key evolution patterns of urban weather systems and then adopted the LSTM network to predict the PM2.5 time series of multiple locations in the city. Zhang et al. [30] proposed an encoder-decoder model for real-time air pollutant prediction, in which LSTM was the main network. The experimental results indicated that the model can fully extract the data correlations and obtain higher prediction accuracy. In addition, the attention mechanism (AM) [31] has attracted extensive attention in time series data analysis and prediction. Han et al. [32] combined LSTM with AM to predict time series, in which the AM can capture time correlation by calculating weights between nodes and neighboring nodes so that it achieved better performance and provided enlightenment for multivariate time series prediction simultaneously.

Although abundant methods have been developed, their performances are limited since the high nonlinearity and nonstationarity of multivariate time series (MTS) data. To improve the prediction performance, a novel encoder-decoder prediction model is presented, and the contributions are as follows:

- (1) The proposed model can sufficiently extract significant temporal features of MTS data.
- (2) As a unit structure, the GRU is adopted to describe sequential characteristics which can reduce model parameters in the procedures of encoding and decoding.
- (3) The AM is introduced into the decoding process for preferably acquiring the reconstructed MTS data.
- (4) To strengthen the prediction performance, 1D-convolution operation and AM are further performed based on the reconstructed new MTS data, which possess discriminant and significant characteristics.

The outline of this paper is as follows. Section 2 reviews the related works, and time series data preprocessing is introduced in Section 3. Section 4 describes the proposed network structure in detail. Section 5 illustrates extensive experiments to verify the effectiveness and feasibility of the proposed model. Section 6 provides some conclusions and future works.

2. Related Works

Recently, researchers have proposed extensive time series (TS) and multivariate time series (MTS) prediction methods, which are classified into two categories including machine learning and deep learning methods [9].

2.1. Machine Learning Methods. The basic assumption of the statistical methods is that the TS and MTS with simple structures are linearity and stationarity. However, in real applications, the TS and MTS data are collected with complex structures, which have high nonlinearity and nonstationarity and they make the TS and MTS forecasting very difficult. Meanwhile, the machine learning algorithms are usually helpful to improve the prediction accuracy [33], which can analyze the behavior of data over time and are independent of the statistical distribution assumption to extract complex nonlinear patterns.

Specifically, Li et al. [34] firstly proposed a chaotic cloud simulated annealing genetic algorithm (CcatCSAGA), which was used to optimize the robust support vector regression (RSVR) parameters for improving the performance of ship traffic flow prediction. Sahoo et al. [35] designed a novel online multiple kernels regression (OMKR), which successively learned kernel-based regression in an extensible manner. Moreover, its effectiveness was demonstrated on real data regression and time series prediction tasks. Ahmed et al. [33], respectively, adopted multilayer perceptron (MLP), Bayesian neural networks (BNN), radial basis function (RBF), general regression neural network (GRNN), k -Nearest neighbors regression (KNNR), classification and regression tree (CART), support vector regression (SVR), and Gaussian process regression (GPR) to perform experiments. This study revealed significant differences between various methods in TS and MTS prediction, and the MLP and GPR methods were the best. Besides, in order to improve the performance, Domingos et al. [36], respectively, combined the ARIMA with MLP and SVR to predict time series. It showed that the hybrid model was better than the single model. Rojas et al. [37] presented a hybrid method integrating an artificial neural network and ARMA model, which achieved outstanding results.

2.2. Deep Learning Methods. The deep neural network can surpassingly learn complex data representation [38], which is widely utilized in many tasks, such as image classification, image segmentation, and natural language processing.

A convolutional neural network (CNN) was originally designed to process static image analysis, which can obtain invariant local relations across spatial dimensions [39]. Recently, CNN and its variant methods were also developed for time series data prediction [40], classification [41], anomaly detection [42], clustering [43], and so on. For example, Ding et al. [44] applied the CNN model to stock market prediction. Wang et al. [45] introduced deep learning to develop a probabilistic wind power generation prediction model. In this model, a wavelet transform was used to decompose the raw wind power data into different frequencies. Then, a CNN model was used to learn nonlinear features in each frequency for improving prediction accuracy. Finally, the probability distribution of wind power generation was predicted. Different from the above methods, Oord et al. [46] proposed a new network model called WaveNet, which expanded convolution to improve the long-term dependence requirement of time series. Moreover, the

size of the receptive field increased exponentially with the depth of layers. Afterward, Borovykh et al. [47] adopted the WaveNet for multivariate financial time series forecasting.

A recurrent neural network (RNN) is also widely exploited for time series prediction [22]. Since there is a long-term dependence on RNN during the training, it will lead to related gradient explosion and gradient disappearance. Therefore, introducing the gating mechanism into RNN has drawn much attention to overcome these limitations and preserves long-term information of time series data, such as long short-term memory (LSTM) [25] and gated recurrent unit (GRU) [26]. The gated variants of RNN essentially preserve the internal state memory through their recurrent feedback mechanism, which makes them very suitable for modeling the time series data. Moreover, their ability to capture complex nonlinear dependence can be extended from short-term to long-term and cross different variables in multivariate systems. Therefore, the performance of these models is excellent in the time series prediction task. Li et al. [48] built a model combining ARIMA and LSTM to improve the prediction accuracy of high-frequency financial time series. Pan et al. [49] applied the model based on the LSTM network to predict urban traffic flow and greatly improved the prediction effect via the spatial correlation. Filonov et al. [50] proposed a model based on the LSTM network to monitor and detect faults in industrial multivariate time series data. Zhao et al. [51] established a two-layer LSTM model to learn gait patterns presenting in neurodegenerative diseases for diagnostic prediction. Jia et al. [52] developed a spatiotemporal learning framework with a dual memory structure based on LSTM to predict land cover. Huang et al. [53] proposed a sequence-to-sequence framework based on GRU to predict different types of abnormal events. Fu et al. [54] used LSTM and GRU to predict short-term traffic flow, which indicated that the RNN-based methods (such as LSTM and GRU) performed better than ARIMA. Zhang et al. [55] utilized four different neural networks, such as MLP, WNN, LSTM, and GRU, to monitor the small watercourses overflow. Furthermore, the models combining CNN with LSTM or GRU have been frequently applied to time series prediction. Wu et al. [56] explored the GRU network to encode the time mode of each sequence with low-dimensional representation and then combined it with a convolutional network for modeling behavioral time series. Shi et al. [57] presented a ConvLSTM network to predict nearby precipitation which can acquire spatiotemporal correlations well.

Autoencoder (AE) has also been successfully applied in time series prediction and is generally combined with other deep learning methods [58]. Considering the inherent temporal and spatial correlation of traffic flow, Lv et al. [59] used AE as one of the modules to construct a deep learning model. Yang et al. [60] proposed a new host load prediction method, which utilized AE as the precyclic feature layer of the echo state network. Gensler et al. [61] combined AE with LSTM for renewable energy power prediction which was superior to the artificial neural network and physical prediction model. Recently, Prenkaj et al. [62] combined AE and GRU to propose a new strategy for predicting the student dropout e-courses.

3. Time Series Data Preprocessing

Generally, time series data are collected manually or automatically; it is difficult to avoid data redundancy, data missing, data error, and other unknown problems in the process of collection and transmission. Therefore, data preprocessing becomes a crucial and necessary procedure for time series data analysis. It mainly includes four stages, such as data clean, data normalization, data sliding window, and data split [63]. The details are illustrated in Figure 1.

- (1) *Data Cleaning*. The purpose of data clean is to deal with missing values, outlier values, and redundant attributes in time series data. There are many ways to handle missing and outlier values. One way is to delete the data with missing and outlier values directly. However, when many attributes of data have missing and outlier values, it is very hard to remain adequate useful attributes and results in incomplete time series data, which will affect the learning and generalization ability of models. The other way considers outlier values as missing values and then the data filling technique is applied to solve the above problems. Data filling includes statistics-based and learning-based methods. The former generally adopts mean filling, while the latter adopts simple linear regression or a complex learning model (such as deep learning). In our work, the mean filling is utilized to process missing values and outlier values. Moreover, feature selection or feature extraction methods are generally adopted to solve redundant attributes. In particular, the proposed model in our work is based on a deep learning framework, which has a strong feature representation ability. Therefore, it is robust to deal with data containing redundant attributes.
- (2) *Data Normalization*. Since the different attributes of data often have different measurement scales, the values collected may vary widely. For the sake of eliminating the influence of measurement scale and value range among different attributes, it is necessary to perform normalized processing which can scale data in a certain proportion, such as mapping data values to $[-1, 1]$ or $[0, 1]$. The popular data normalization methods contain minimum-maximum normalization and zero-mean normalization.

Minimum-maximum normalization is named deviation standardization, which maps the values of the original data to $[0, 1]$ via a linear transformation. The formula is as follows:

$$x^* = \frac{x - \min}{\max - \min}, \quad (1)$$

where \max and \min represent the maximum and minimum values of data, respectively. The method can preserve the relationships that exist in original data.

Zero-mean normalization is known as standard deviation standardization. After processing, the

mean value and the standard deviation of normalization data are 0 and 1, respectively. The formula is defined as

$$x^* = \frac{x - \bar{x}}{\sigma}, \quad (2)$$

where \bar{x} and σ are the mean and standard deviation of original data, respectively.

- (3) *Data Sliding Window*. This operation mainly creates time series data by the predefined sliding window size and step for the original time series data. In other words, this operation is used to generate the predicted data for the next moment using historical data with a given interval. The specific operation of the data sliding window is shown in Figure 2 [64]. Given any time series data with length N , such as $\{1, 2, 3, 4, 5, \dots, N-1, N\}$, when the sliding window size is set to L and the sliding step is 1, the $N-L$ data sets with length $L+1$ are formed. Particularly, the first L data of each set is regarded as training data and the value of the number $L+1$ is the target value.
- (4) *Data Split*. This stage divides the time series dataset into training data and test data. For example, the first 60% are used for training and the remaining 40% are used to test in the experiments.

4. The Proposed Method

In this work, a novel time series prediction model based on the encoding-decoding framework is designed, which integrates the recurrent neural module, convolutional module, attention mechanism, and fully connection module into a unified framework. As shown in Figure 3, the proposed model consists of three parts such as encoding, decoding, and prediction modules. In the encoding module, the gated recurrent unit (GRU) is taken as the main unit structure for extracting more effective time series features. In the decoding module, the attention mechanism (AM) is introduced to explore the importance of historical data collected at different times, so that it can obtain better new time series data. In addition, taking the influence of previous features on data reconstruction into account, feature reuse is realized by the skip connections based on the residual network. In the prediction module, the convolution layer is adopted to extract effective features from the reconstruction time series. Then, the AM is further performed on the convolution feature mapping owing to the influence of important information on prediction performance. Finally, a multilayer fully connected network is established for prediction.

4.1. Deep Autoencoder (DAE). Autoencoder (AE) is an unsupervised deep learning method which is frequently used in feature representation, data compression, image denoising, and other tasks [28]. The structure of AE includes an encoder and a decoder, which only contain a fully connected hidden layer. To better extract features and reconstruct original data, Deep Autoencoder (DAE) [65] is

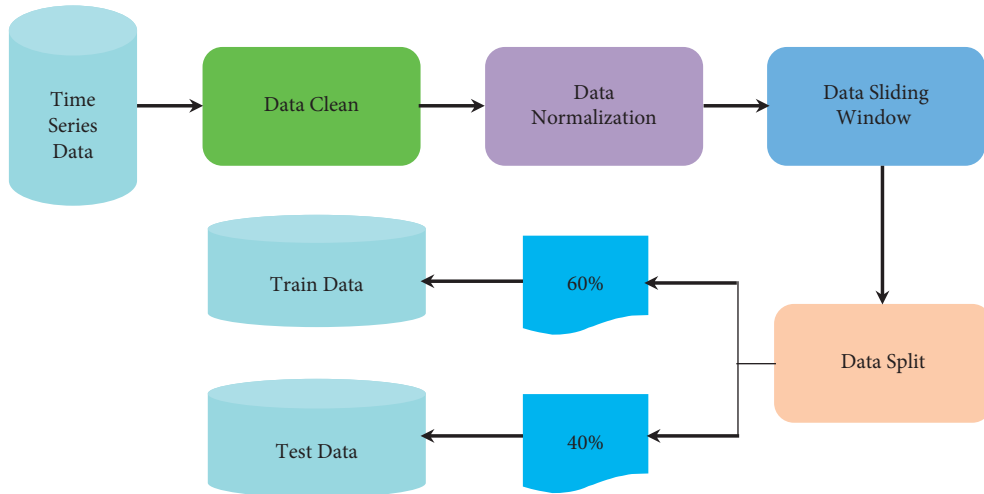


FIGURE 1: The process of time series data preprocessing.

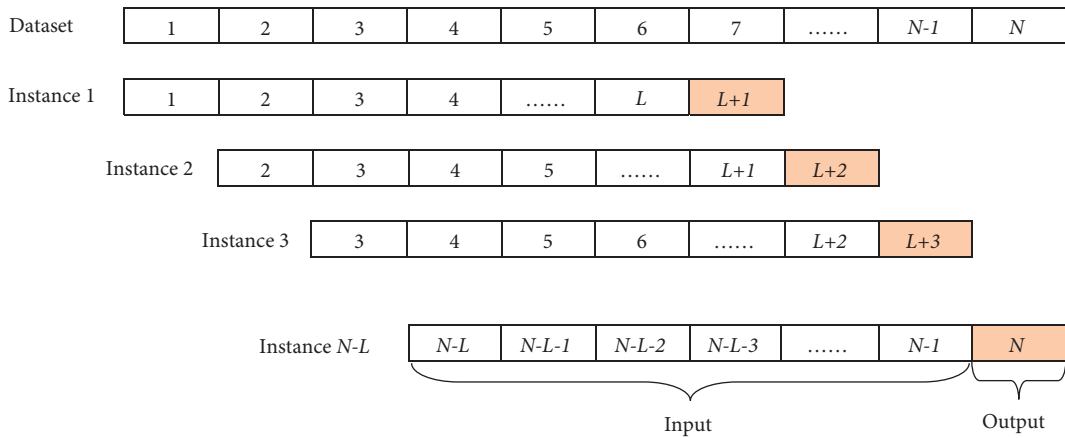


FIGURE 2: The process of data sliding window for creating a time series data.

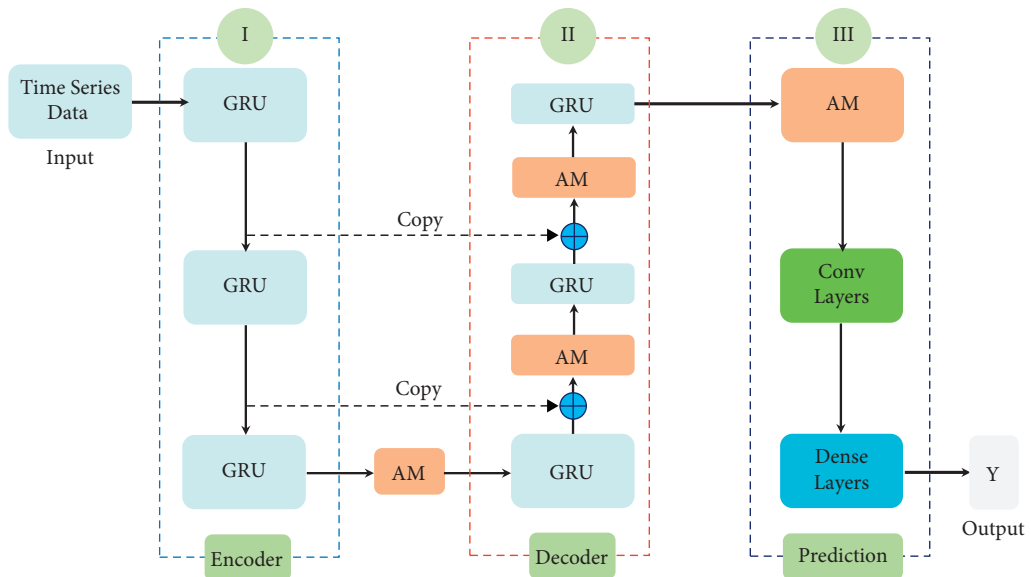


FIGURE 3: The structure of the proposed network model.

designed that contains multiple hidden layers shown in Figure 4.

4.2. LSTM and GRU. In general, DAE is a multilayer feedforward neural network, while it does not consider the importance of historical information of time series data to the prediction or classification of unknown data. As a specific network structure, a recurrent neural network (RNN) [22] can adeptly utilize the historical information of time series data, which adopts a backpropagation through time (BPTT) algorithm to train and learn parameters. However, RNN produced gradient vanishing or gradient exploding problems when it handled time series with long time intervals [25]. In particular, the longer the time interval, the more likely it is to appear severe gradient vanishing or gradient exploding, which will make it difficult to train effective RNN models for long interval sequences.

To solve the above problems, other RNN variants (such as LSTM [25] and GRU [26]) are easier to capture the long-term dependence of time series data. LSTM uses the gate mechanism to control the information accumulation speed and can selectively update information and forget information accumulated. LSTM includes an input gate, forget gate, and output gate, which are displayed in Figure 5. The forget gate f_t controls which information needs to be forgotten derived from the internal state of the previous moment. The input gate i_t controls which information from the current candidate state needs to be retained. And the output gate o_t controls which information of the current internal state needs to be output.

Different from LSTM, GRU is a simplified version of LSTM. It merges the forget gate and input gate into the update gate and retains the original reset gate, as shown in Figure 6. It can be observed that no additional memory units are needed in GRU. It is due to the fact that an update gate can control how much information needs to retain from the historical state and needs to receive from the candidate state for the current state. The calculation formula of GRU is

$$\begin{aligned} z_t &= \sigma(W_z x_t + U_z h_{t-1} + b_z), \\ r_t &= \sigma(W_r x_t + U_r h_{t-1} + b_r), \\ \tilde{h}_t &= \tanh(W_h x_t + r_t \odot U_h h_{t-1} + b_h), \\ h_t &= z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t, \end{aligned} \quad (3)$$

where z_t and r_t represent update gate and reset gate, respectively. h_t is the state of the current moment t and \tilde{h} indicates the candidate state. σ is the sigmoid activation function that can convert results to $[0, 1]$. \tanh stands for hyperbolic tangent activation function. The symbol \odot is the dot product operation of corresponding elements. x_t represents the input of the neural network at time t . W_z , W_r , W_h and U_z , U_r , U_h represent the parameter matrix and recurrent weight of the model. b_z , b_r , and b_h are the offset vector. Compared with LSTM, GRU has a simple structure and fewer parameters because there are fewer gate structures of GRU. Therefore, GRU not only can reduce the model training time and avoid overfitting problems but also can

achieve the same results as LSTM and even better than LSTM. In addition, BiGRU is a variant version of GRU. Although BiGRU has better performance than GRU in some cases, the parameter size of BiGRU is bigger than GRU. In order to overcome the overfitting problem, the GRU is adopted as the main unit structure of the autoencoder.

4.3. Attention Mechanism. Attention mechanism (AM) has been widely applied to natural language, computer vision, and other fields [66]. It is a resource allocation scheme that uses limited computing resources to process more important information for the information overload problem. Like artificial neural networks, AM originated from human vision and borrowed from human visual attention mechanisms. The core idea of AM is to select the more critical information and ignore the unimportant or irrelevant information to the current task from a large amount of information [66]. At present, plenty of attention mechanisms have been built to solve related tasks, such as spatial attention, channel attention, and mixed attention mechanisms [67].

In image understanding tasks including image segmentation and target detection, the channel attention (CA) [68] module is mainly adopted to explore relationships between feature maps of different channels, and its structure is shown in Figure 7. In the module, the feature map of each channel is taken as a feature detector that can determine which part of the features should be noticed more. It is well known that the time attribute is very important and also affects the prediction results. Therefore, we view each time attribute as a channel and the channel attention (CA) mechanism is integrated to mine the significance of time attributes in the proposed method.

4.4. Prediction Module. In the prediction module, a 1D-convolution is firstly explored to extract features from the time series data reconstructed by DAE. Then, in order to explore the different contributions of historical data for forecasting, the CA mechanism is performed on feature mapping by the previous layer. Finally, a multilayer dense network structure is constructed for prediction. The details are displayed in Figure 8.

5. Experiments and Results Analysis

To verify the effectiveness of the proposed method, two series of experiments are conducted on public stock and shared bicycle datasets, respectively, and compared with some related methods. Many experimental results validate the effectiveness of our model.

5.1. Evaluation Metrics and Experimental Environment. In order to quantitatively analyze the accuracy and superiority, mean square error (MSE), mean square error (RMSE), mean absolute error (MAE), and mean average percentage error (MAPE) are adopted to evaluate the performance of the proposed model [69]. The calculation formulas are as follows:

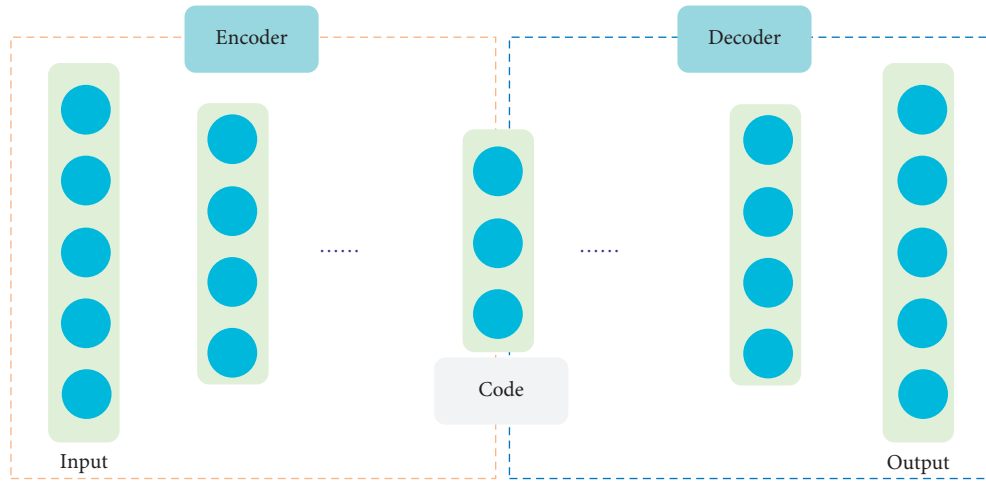


FIGURE 4: The structure of DAE.

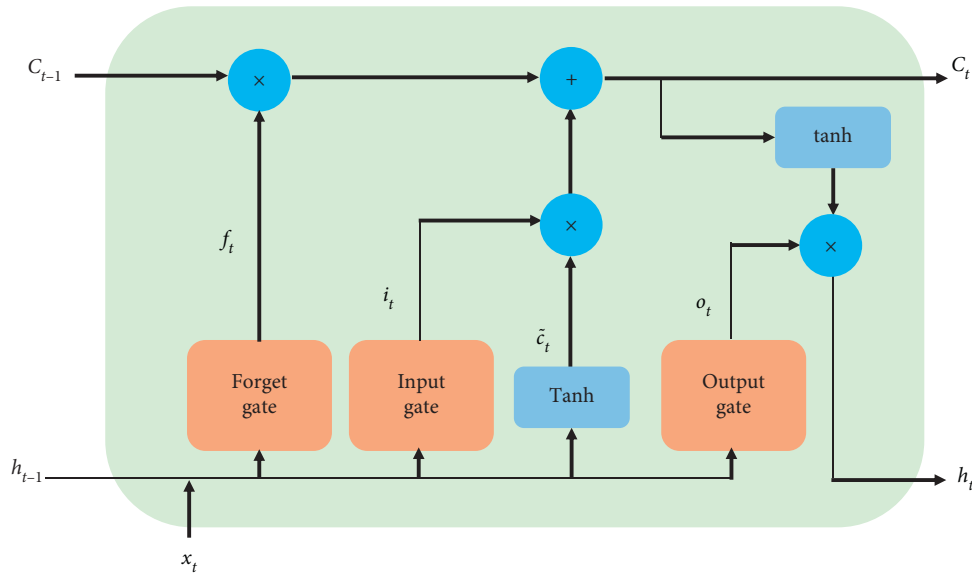


FIGURE 5: The structure of the LSTM unit.

$$\begin{aligned}
 \text{MSE} &= \frac{1}{n} \sum_{t=1}^n (X_t - X'_t)^2, \\
 \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{t=1}^n (X_t - X'_t)^2}, \\
 \text{MAE} &= \frac{1}{n} \sum_{t=1}^n |X_t - X'_t|, \\
 \text{MAPE} &= \frac{100}{n} \sum_{t=1}^n \frac{|X_t - X'_t|}{|X'_t|},
 \end{aligned} \tag{4}$$

where X_t and X'_t represent the actual and predicted values of the data and n is the number of samples. The smaller the above values, the more accurate the prediction result.

The source codes of the proposed method and the compared methods are completed using Tensorflow with Python. The corresponding versions of the development software and the configurations of the hardware platform are listed in Table 1. Moreover, the settings of the key parameters during the training processing are shown in Table 2.

5.2. Stock Data Prediction

5.2.1. Stock Data Description. The stock data used in the experiment are Shanghai Composite Index 50 (SCI-50), CSI-300, and Shenzhen Component Index (SZCI). Each stock data records multiple attributes, such as the closing price, the highest price, the lowest price, the opening price, the previous day's closing price, change, and ups and downs. The closing price, the highest price, the lowest price, and the opening price represent the final price, the highest price, the

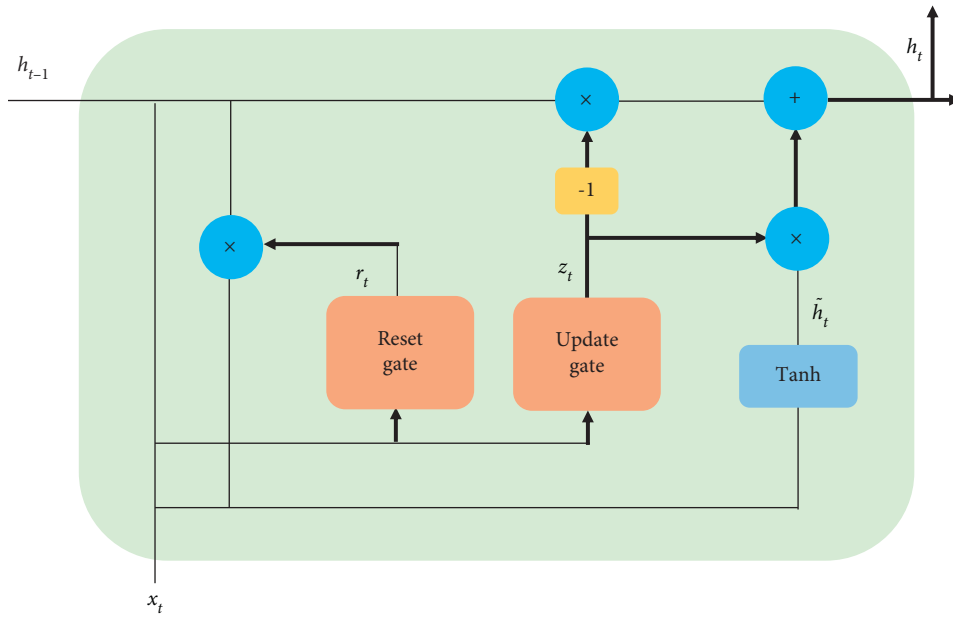


FIGURE 6: The structure of the GRU unit.

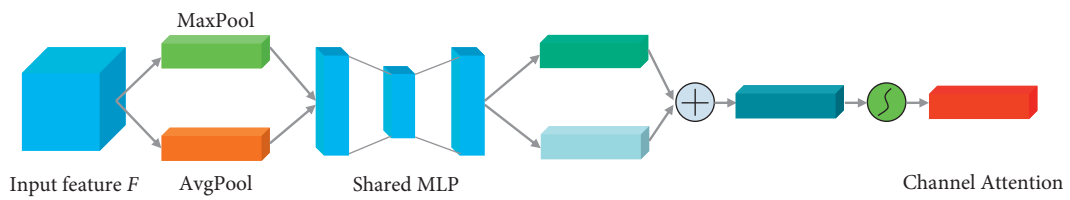


FIGURE 7: The structure of CA.

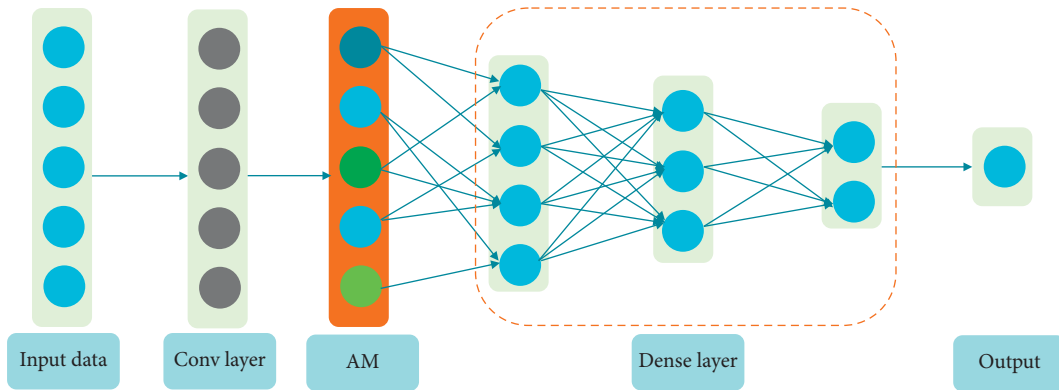


FIGURE 8: The structure of the prediction module.

TABLE 1: The description of experimental environment.

<i>Development software</i>	<i>Version</i>
Python	3.6.0
Tensorflow	2.7.0
System	Window 10 64 bit
<i>Hardware platform</i>	<i>Configurations</i>
PC machine	Inter core i9 9900k
RAM	32 GB
GPU	GeForce RTX 2080 Ti GPU

TABLE 2: The settings of the key parameters in the training procedure.

Description	Value
Batch-size	256
Optimizer	Adam
Epochs	400
Loss function	MSE

lowest price, and the first trading price of one stock, respectively. The previous day’s closing price is the final price at which a stock is traded on the previous day. Change is the difference between the closing price and the previous day’s closing price of the stocks traded (i.e., closing price - previous day’s closing price). The value of ups and downs is the change divided by the closing price of the stocks traded (i.e., change/closing price). The details of three stock datasets are listed in Table 3. Meanwhile, Tables 4 to 6 give some instances of data and corresponding statistical information for each stock, including the number of records, minimum, maximum, mean, variance, 1/4 value, 1/2 value, and 3/4 value for each attribute. From Tables 3 to 5, we can see that there are great differences and fluctuations in the stock data.

5.2.2. Parameters Analysis. Time interval (time step) is the significant factor affecting the prediction of time series data. Therefore, we test the performance of the proposed method with different steps. In the experiment, the time step is set to {5, 10, 15, 20, 25, 30}, and the experimental results are displayed in Tables 7–9. Obviously, in most cases, when the step increases, the value of each evaluation indicator decreases. It indicates that the performance of the proposed model improves with the increasing step. This is because long interval data provides more useful information for prediction. However, as the step continues to increase, the values of each evaluation indicator will increase. It indicates that the performance of the proposed model decreases with the increase of time step. The possible reason is that time series data with too long intervals contains redundant information and high volatility, which makes it difficult to capture more effective information for future data prediction.

5.2.3. Convergence Analysis. In order to verify the convergence of our proposed method, we plot the curves of loss values (MSE) on the training set and validation set for each dataset. From Figure 9, we can see that our model reaches convergent very quickly on the training set. For the validation set, the loss values (MSE) of the proposed model fluctuate but basically maintains stability when the number of iteration (Epochs) is greater than 400.

5.2.4. Performance Analysis. In order to further test the performance of the proposed method, we compare it with GRU, BiGRU, GRU-AE, BiGRU-AE, GRU-AE-AM, and BiGRU-AE-AM. Tables 10–12 show the results of different methods on three stock datasets. The following conclusions can be drawn from the experimental results:

TABLE 3: The details of three stock datasets.

Stock name	Stock code	Start and end time	Number of records
SCI-50	000016	2004.01.02–2021.06.23	4245
CSI-300	399300	2002.01.07–2021.03.17	4657
SZCI	399001	1991.04.04–2021.06.23	7349

- (1) The performances of traditional GRU and BiGRU models are lower than those of other comparison methods. Furthermore, BiGRU not only makes use of the useful information of historical data in the forward direction but also mines the dependence of current data on historical data in the reverse direction. Therefore, BiGRU has better performance than GRU.
- (2) The performances of recurrent neural networks (GRU-AE and BiGRU-AE) are superior to the traditional recurrent neural network (GRU and BiGRU). It indicates that introducing encoding-decoding into the recurrent neural network is beneficial to improving the prediction performance of the proposed model.
- (3) The performances of the recurrent neural network-AE model based on attention mechanisms (GRU-AE-AM and BiGRU-AE-AM) exceed the recurrent neural network-AE model (GRU-AE and BiGRU-AE). It demonstrates that introducing the attention mechanism into the recurrent neural network can mine significant information in time series data.
- (4) The proposed model is based on the idea of integrating encoding-decoding and attention mechanisms simultaneously into the recurrent neural network. Different from GRU-AE-AM and BiGRU-AE-AM, the proposed method develops the attention mechanism in the decoding stage to capture the degree of importance between different intervals. Therefore, compared with other methods, the presented method establishes significant advantages on different evaluation indicators.

5.3. Demand Forecast for Shared Bicycle Data

5.3.1. Shared Bicycle Data Description. The datasets of this experiment are derived from the shared bicycle demand of three streets in Shenzhen, China, such as Longgang Central City, Pingshan Street, and Zhaoshang Street. Each data set contains the historical travel data of shared bicycles, time attribute data (such as hours, working day or not), and weather data (such as temperature, rainfall, wind speed, and humidity). The details are listed in Table 13.

5.3.2. Parameters Analysis. In this experiment, the influence of the time step on the prediction performance is also analyzed adequately. The step size setting is consistent with stock price prediction experiments, and experimental results are shown in Tables 14–16. We can see that the effect

TABLE 6: Continued.

Data	Closing price	Highest price	Lowest price	Opening price	Previous day's closing price	Change	Ups and downs
Mean	6709.184	6778.63	6628.694	6704.283	6707.301	1.885397	0.05939
Std	4325.842	4369.826	4270.334	4322.335	4325.313	153.7217	2.1302
Min	402.5	408.02	397.67	401.57	402.5	-1293.66	-19.7807
25%	3112.336	3134.055	3077.097	3112.637	3111.4	-42.702	-0.8978
50%	4834.614	4867.142	4795.043	4836.637	4831.989	0.381	0.0112
75%	10316.82	10410.65	10223.16	10315	10315.75	51.813	0.9835
Max	19531.16	19600.03	19203.11	19554.58	19531.16	1254.795	26.1963

TABLE 7: The results with different steps on SCI-50.

Time step	MSE	RMSE	MAE	MAPE
5	1682.935	41.024	27.188	1.023
10	1673.594	40.910	27.266	1.025
15	1736.988	41.677	27.588	1.036
20	1757.061	41.917	28.304	1.062
25	1752.673	41.865	28.084	1.055
30	1780.636	42.198	28.547	1.072

Bold in the table indicates the optimal results.

TABLE 8: The results with different steps on CSI-300.

Time step	MSE	RMSE	MAE	MAPE
5	3157.709	56.193	36.205	1.001
10	3085.284	55.545	36.214	1.005
15	3233.284	56.862	36.904	1.025
20	3287.964	57.341	37.026	1.026
25	3438.429	58.638	39.390	1.082
30	3393.111	58.250	38.940	1.069

Bold in the table indicates the optimal results.

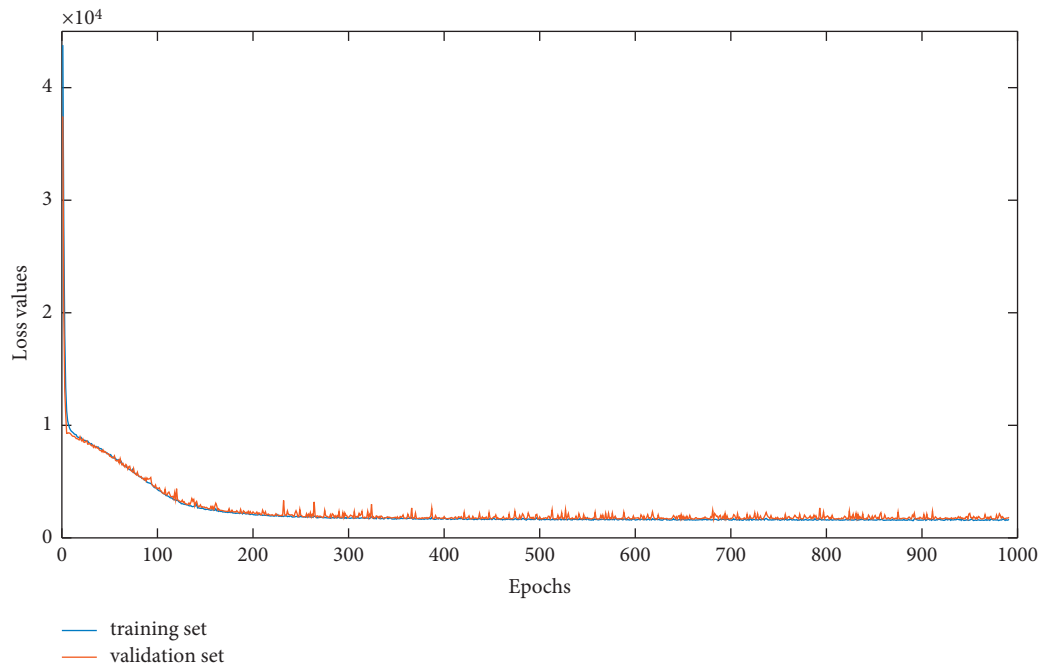
TABLE 9: The results with different steps on SZCI.

Time step	MSE	RMSE	MAE	MAPE
5	34522.267	185.802	127.234	1.186
10	33851.601	183.988	127.063	1.180
15	34495.000	185.730	128.299	1.190
20	34767.899	186.462	128.943	1.195
25	36065.960	189.910	132.394	1.226
30	36287.302	190.492	132.812	1.228

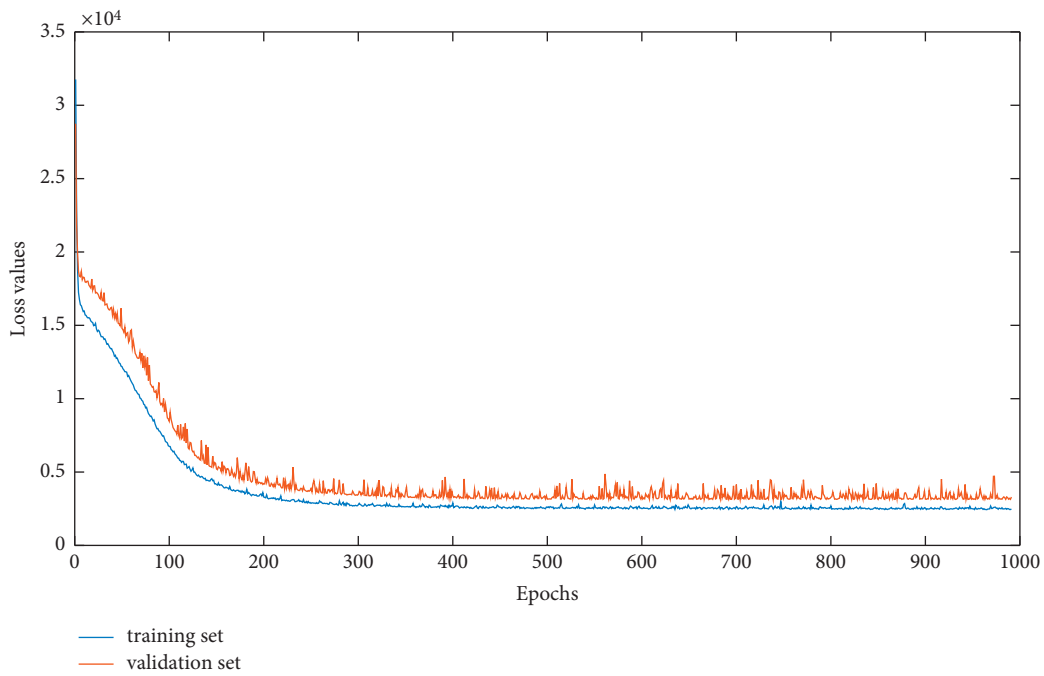
Bold in the table indicates the optimal results.

of steps differs from the experimental results on the stock data. Firstly, with the increase of steps, the evaluation indicator values of the proposed method decreased on Longgang and Pingshan datasets. However, this trend does not always remain unchanged, and the opposite result will occur when the step continues to increase. Accordingly, the performance of the proposed model will also decrease. Secondly, the results are different from Tables 7–9 and Tables 14–16 on the Longgang Street dataset. When the step is set to the minimum ($L=5$), the proposed method can obtain the optimal results. The cause is maybe that time series data has strong dependence and complex data structure.

5.3.3. *Performance Analysis.* Similarly, the proposed method is compared with other well-known methods, and the results are shown in Tables 17–19. On the whole, the experimental results are consistent with those of stock experiments, except for the data in Longgang. In particular, the proposed method can achieve better performance with a step value of 20. It indicates that the data structure is relatively simple, which is prone to overfitting for the complex model. Therefore, the evaluation metrics of the bidirectional recurrent neural network model (BiGRU, BiGRU-AE, and BiGRU-AE-AM) are higher than those of the recurrent neural network model with unidirectional structure (GRU, GRU-AE, and GRU-AE-AM).

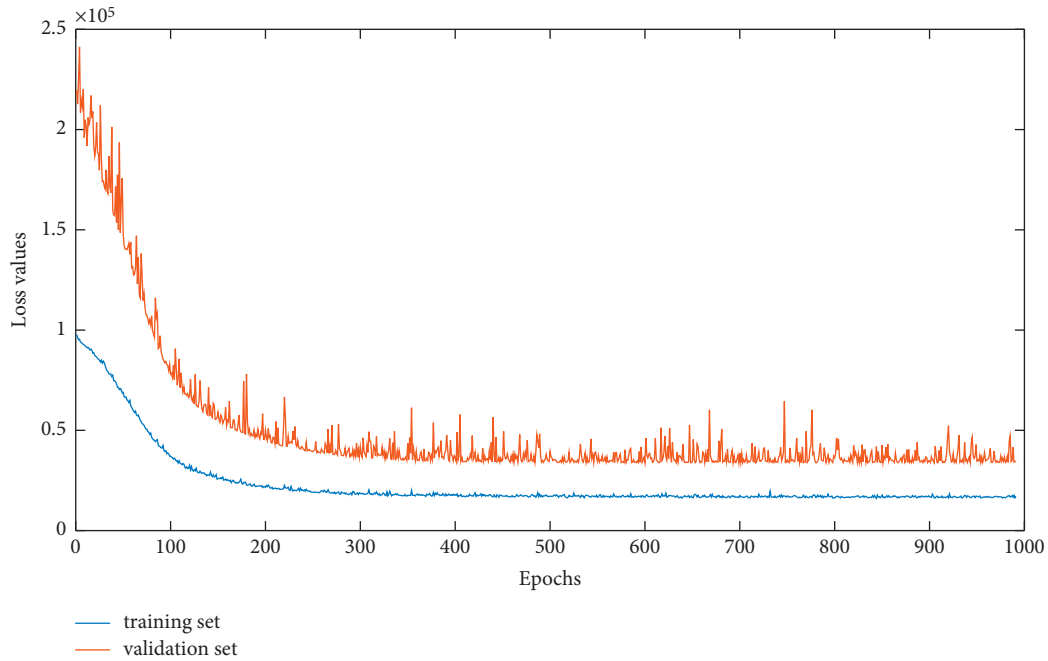


(a)



(b)

FIGURE 9: Continued.



(c)

FIGURE 9: The curves of loss values (MSE) on the training set and validation set of three stock datasets. (a) SCI-50. (b) CSI-300. (c) SZCI.

TABLE 10: The results with step value of 10 on SCI-50.

Method	MSE	RMSE	MAE	MAPE
GRU	2356.925	48.548	35.958	1.328
BiGRU	2267.462	47.618	35.466	1.342
GRU-AE	2371.064	48.694	37.074	1.419
BiGRU-AE	1964.477	44.322	31.129	1.164
GRU-AE-AM	2040.477	45.172	31.334	1.164
BiGRU-AE-AM	1814.952	42.602	28.483	1.062
Our method	1673.594	40.910	27.266	1.025

Bold in the table indicates the optimal results.

TABLE 11: The results with step value of 10 on CSI-300.

Method	MSE	RMSE	MAE	MAPE
GRU	4262.664	65.289	46.799	1.264
BiGRU	3614.219	60.118	41.625	1.137
GRU-AE	3382.457	58.159	38.588	1.070
BiGRU-AE	3828.393	61.874	44.642	1.244
GRU-AE-AM	3798.575	61.633	41.392	1.127
BiGRU-AE-AM	3726.034	61.041	39.880	1.084
Our method	3085.284	55.545	36.214	1.005

Bold in the table indicates the optimal results.

TABLE 12: The results with step value of 10 on SZCI.

Method	MSE	RMSE	MAE	MAPE
GRU	37269.796	193.054	137.383	1.271
BiGRU	35012.771	187.114	126.924	1.180
GRU-AE	34737.291	186.379	128.821	1.203
BiGRU-AE	37163.139	192.777	134.793	1.257
GRU-AE-AM	35198.463	187.613	130.611	1.218
BiGRU-AE-AM	34946.619	186.940	131.015	1.216
Our method	33851.601	183.988	127.063	1.180

Bold in the table indicates the optimal results.

TABLE 13: The description of shared bicycle datasets.

Dataset	Time	Quantity by hour
Longgang central city	2016.6–2017.8 (except Dec.)	6935
Pingshan street	2016.7–2017.8 (except Dec.)	6935
Zhaoshang street	2016.7–2016.11	2907

TABLE 14: The results with different steps of shared bicycle data on Longgang.

Time step	MSE	RMSE	MAE	MAPE
5	684.764	26.168	17.429	102.576
10	663.629	25.761	16.881	89.556
15	672.780	25.938	16.598	79.102
20	652.445	25.543	16.421	88.784
25	726.195	26.948	17.697	93.377
30	695.377	26.370	17.800	123.276

Bold in the table indicates the optimal results.

TABLE 15: The results with different steps of shared bicycle data on Pingshan.

Time step	MSE	RMSE	MAE	MAPE
5	240.870	15.520	11.778	20.356
10	227.618	15.087	11.386	17.991
15	222.815	14.927	11.247	17.931
20	238.981	15.459	12.046	22.808
25	228.705	15.123	11.449	19.670
30	224.910	14.997	11.343	17.497

Bold in the table indicates the optimal results.

TABLE 16: The results with different steps of shared bicycle data on Zhaoshang.

Time step	MSE	RMSE	MAE	MAPE
5	1071.253	32.730	22.051	76.338
10	1084.648	32.934	22.286	58.965
15	1282.643	35.814	24.423	63.207
20	1201.246	34.659	23.586	63.110
25	1322.340	36.364	24.477	62.043
30	1430.125	37.817	25.936	67.414

Bold in the table indicates the optimal results.

TABLE 17: The results with step 20 of shared bicycle data on Longgang.

Method	MSE	RMSE	MAE	MAPE
GRU	717.634	26.789	17.791	76.668
BiGRU	718.049	26.796	17.711	63.458
GRU-AE	784.713	28.013	18.660	106.267
BiGRU-AE	904.714	30.078	22.273	163.759
GRU-AE-AM	740.382	27.210	17.726	85.522
BiGRU-AE-AM	828.928	28.791	18.109	91.725
Our method	652.445	25.543	16.421	88.784

Bold in the table indicates the optimal results.

TABLE 18: The results with step 15 of shared bicycle data on Pingshan.

Method	MSE	RMSE	MAE	MAPE
GRU	308.121	17.553	13.758	21.750
BiGRU	229.627	15.153	11.481	17.345
GRU-AE	275.273	16.591	12.690	19.307
BiGRU-AE	270.095	16.435	12.397	17.419

TABLE 18: Continued.

Method	MSE	RMSE	MAE	MAPE
GRU-AE-AM	212.592	14.581	10.800	15.397
BiGRU-AE-AM	211.903	14.557	10.876	15.818
Our method	222.815	14.927	11.247	17.931

Bold in the table indicates the optimal results.

TABLE 19: The results with step 5 of shared bicycle data on Zhaoshang.

Method	MSE	RMSE	MAE	MAPE
GRU	1174.327	34.268	24.864	55.381
BiGRU	1139.840	33.762	24.263	55.748
GRU-AE	1124.142	33.528	24.134	86.665
BiGRU-AE	1180.541	34.359	23.882	72.879
GRU-AE-AM	1241.616	35.237	23.773	46.250
BiGRU-AE-AM	1195.643	34.578	22.705	43.788
Our method	1071.253	32.730	22.051	76.338

Bold in the table indicates the optimal results.

6. Conclusions and Future Works

In this paper, to improve the accuracy of time series data prediction, the autoencoder, recurrent neural network, attention mechanism, convolution module, and full connection module are integrated to establish a novel prediction model based on an encoding-decoding framework. The prediction performances are evaluated for the stock price and the demand for shared bicycles on three stock datasets and three shared bicycle datasets, respectively. In addition, we compare it with many other related methods, which demonstrate that the proposed model has higher prediction accuracy from the views of multiple quantitative indicators (such as MSE, RMSE, MAE, and MAPE).

The future works mainly include the following points. (1) We will try to apply the proposed model to prediction tasks of time series data in other fields (such as medical, energy, environment, and other industrial data). (2) Using the core idea, we will further extend it to solve the anomaly detection task of time series data. (3) We will intensively study how to combine the traditional multivariate time series method with deep learning to further improve the prediction performance in real applications.

Data Availability

The network code and data are available from the corresponding author upon request.

Conflicts of Interest

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

Acknowledgments

This work was supported by the Shandong Provincial Natural Science Foundation (No. ZR2021MF026), the Doctoral Program of Weifang University (2018BS11), the Science and Technology Program of Public Wellbeing (Nos.

2020KJHM01 and 2021KJHM08), the National Natural Science Foundation of China (Nos. 62062040, 62006174, 61832012, 61672321, and 61771289), and the National Key R&D Program of China (No. 2019YFB2102600).

References

- [1] A. Blázquez-García, A. Conde, U. Mori, and J. A. Lozano, "A review on outlier/anomaly detection in time series data," *ACM Computing Surveys*, vol. 54, no. 3, pp. 1–33, 2021.
- [2] J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, and A. Troncoso, "Deep learning for time series forecasting: a survey," *Big Data*, vol. 9, no. 1, pp. 3–21, 2021.
- [3] P. Lara-Benítez, M. Carranza-García, and J. C. Riquelme, "An experimental review on deep learning architectures for time series forecasting," *International Journal of Neural Systems*, vol. 31, no. 3, Article ID 2130001, 2021.
- [4] Z. Han, J. Zhao, H. Leung, K. Fai Ma, and W. Wang, "A review of deep learning models for time series prediction," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7833–7848, 2019.
- [5] K. Yadav, M. Yadav, and S. Saini, "Stock values predictions using deep learning based hybrid models," *CAAI Transactions on Intelligence Technology*, vol. 7, no. 1, pp. 107–116, 2022.
- [6] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8387680, 11 pages, 2021.
- [7] M. Chen, X. Yu, and Y. Liu, "PCNN: deep convolutional networks for short-term traffic congestion prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 11, pp. 3550–3559, 2018.
- [8] S. Du, T. Li, Y. Yang, and S. J. Horng, "Deep air quality forecasting using hybrid deep learning framework," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 6, pp. 2412–2424, 2019.
- [9] Z. Liu, Z. Zhu, J. Gao, and C. Xu, "Forecast methods for time series data: a survey," *IEEE Access*, vol. 9, pp. 91896–91912, 2021.
- [10] 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.
- [11] S. Dong, P. Wang, and K. Abbas, “A survey on deep learning and its applications,” *Computer Science Review*, vol. 40, Article ID 100379, 2021.
 - [12] K. S. Tuncel and M. G. Baydogan, “Autoregressive forests for multivariate time series modeling,” *Pattern Recognition*, vol. 73, pp. 202–215, 2018.
 - [13] H. Yang, Z. Pan, Q. Tao, and J. Qiu, “Online learning for vector autoregressive moving-average time series prediction,” *Neurocomputing*, vol. 315, pp. 9–17, 2018.
 - [14] Z.-X. Wang, Y.-F. Zhao, and L.-Y. He, “Forecasting the monthly iron ore import of China using a model combining empirical mode decomposition, non-linear autoregressive neural network, and autoregressive integrated moving average,” *Applied Soft Computing*, vol. 94, Article ID 106475, 2020.
 - [15] N. Sapankevych and R. Sankar, “Time series prediction using support vector machines: a survey,” *IEEE Computational Intelligence Magazine*, vol. 4, no. 2, pp. 24–38, 2009.
 - [16] B. Dong, C. Cao, and S. E. Lee, “Applying support vector machines to predict building energy consumption in tropical region,” *Energy and Buildings*, vol. 37, no. 5, pp. 545–553, 2005.
 - [17] R. N. Yadav, P. K. Kalra, and J. John, “Time series prediction with single multiplicative neuron model,” *Applied Soft Computing*, vol. 7, no. 4, pp. 1157–1163, 2007.
 - [18] D. K. Barrow and S. F. Crone, “A comparison of AdaBoost algorithms for time series forecast combination,” *International Journal of Forecasting*, vol. 32, no. 4, pp. 1103–1119, 2016.
 - [19] M. Ghiassi, H. Saidane, and D. K. Zimbra, “A dynamic artificial neural network model for forecasting time series events,” *International Journal of Forecasting*, vol. 21, no. 2, pp. 341–362, 2005.
 - [20] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
 - [21] K. Wang, K. Li, L. Zhou et al., “Multiple convolutional neural networks for multivariate time series prediction,” *Neurocomputing*, vol. 360, pp. 107–119, 2019.
 - [22] Y. Liu, C. Gong, L. Yang, and Y. Chen, “DSTP-RNN: a dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction,” *Expert Systems with Applications*, vol. 143, Article ID 113082, 2020.
 - [23] J. Zhang, Y. Zheng, D. Qi, R. Li, X. Yi, and T. Li, “Predicting citywide crowd flows using deep spatio-temporal residual networks,” *Artificial Intelligence*, vol. 259, pp. 147–166, 2018.
 - [24] A. N. Jagannatha and H. Yu, “Bidirectional RNN for medical event detection in electronic health records,” in *Proceedings of the conference. Association for Computational Linguistics. North American Chapter. Meeting*, pp. 473–482, San Diego, California, June 2016.
 - [25] 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.
 - [26] C. Li, G. Tang, X. Xue, A. Saeed, and X. Hu, “Short-term wind speed interval prediction based on ensemble GRU model,” *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1370–1380, 2019.
 - [27] W. Shu, K. Cai, and N. N. Xiong, “A short-term traffic flow prediction model based on an improved gate recurrent unit neural network,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.
 - [28] C. Zhang, Y. Liu, and H. Fu, “Ae2-nets: autoencoder in autoencoder networks,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2577–2585, Long Beach, CA, USA, June 2019.
 - [29] X. Xu and M. Yoneda, “Multitask air-quality prediction based on LSTM-autoencoder model,” *IEEE Transactions on Cybernetics*, vol. 51, no. 5, pp. 2577–2586, 2021.
 - [30] B. Zhang, G. Zou, D. Qin, Y. Lu, Y. Jin, and H. Wang, “A novel Encoder-Decoder model based on read-first LSTM for air pollutant prediction,” *The Science of the Total Environment*, vol. 765, Article ID 144507, 2021.
 - [31] S. Du, T. Li, Y. Yang, and S.-J. Horng, “Multivariate time series forecasting via attention-based encoder-decoder framework,” *Neurocomputing*, vol. 388, pp. 269–279, 2020.
 - [32] S. Han, H. Dong, X. Teng, X. Li, and X. Wang, “Correlational graph attention-based Long Short-Term Memory network for multivariate time series prediction,” *Applied Soft Computing*, vol. 106, Article ID 107377, 2021.
 - [33] N. K. Ahmed, A. F. Atiya, N. E. Gayar, and H. El-Shishiny, “An empirical comparison of machine learning models for time series forecasting,” *Econometric Reviews*, vol. 29, no. 5-6, pp. 594–621, 2010.
 - [34] M.-W. Li, D.-F. Han, and W. L. Wang, “Vessel traffic flow forecasting by RSVR with chaotic cloud simulated annealing genetic algorithm and KPCA,” *Neurocomputing*, vol. 157, pp. 243–255, 2015.
 - [35] D. Sahoo, S. C. H. Hoi, and B. Li, “Large scale online multiple kernel regression with application to time-series prediction,” *ACM Transactions on Knowledge Discovery from Data*, vol. 13, no. 1, pp. 1–33, 2019.
 - [36] S. O. Domingos, J. F. L. de Oliveira, and P. S. G. de Mattos Neto, “An intelligent hybridization of ARIMA with machine learning models for time series forecasting,” *Knowledge-Based Systems*, vol. 175, pp. 72–86, 2019.
 - [37] I. Rojas, O. Valenzuela, F. Rojas et al., “Soft-computing techniques and ARMA model for time series prediction,” *Neurocomputing*, vol. 71, no. 4-6, pp. 519–537, 2008.
 - [38] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, “A survey of deep neural network architectures and their applications,” *Neurocomputing*, vol. 234, pp. 11–26, 2017.
 - [39] W. Rawat and Z. Wang, “Deep convolutional neural networks for image classification: a comprehensive review,” *Neural Computation*, vol. 29, no. 9, pp. 2352–2449, 2017.
 - [40] S. Liu, H. Ji, and M. C. Wang, “Nonpooling convolutional neural network forecasting for seasonal time series with trends,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 8, pp. 2879–2888, 2019.
 - [41] C. L. Liu, W. H. Hsaio, and Y. C. Tu, “Time series classification with multivariate convolutional neural network,” *IEEE Transactions on Industrial Electronics*, vol. 66, no. 6, pp. 4788–4797, 2018.
 - [42] N. Chouhan, A. Khan, and H. U. R. Khan, “Network anomaly detection using channel boosted and residual learning based deep convolutional neural network,” *Applied Soft Computing*, vol. 83, Article ID 105612, 2019.
 - [43] O. F. Özgül, B. Bardak, and M. Tan, “A convolutional deep clustering framework for gene expression time series,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 6, pp. 2198–2207, 2020.
 - [44] X. Ding, Y. Zhang, T. Liu, and J. Duan, “Deep Learning for Event-Driven Stock prediction,” in *Proceedings of the Twenty-fourth International Joint Conference on Artificial Intelligence, AAAI, Buenos Aires, Argentina, July 2015*.

- [45] H.-z. Wang, G.-q. Li, G.-b. Wang, J.-c. Peng, H. Jiang, and Y.-t. Liu, "Deep learning based ensemble approach for probabilistic wind power forecasting," *Applied Energy*, vol. 188, pp. 56–70, 2017.
- [46] A. Oord, S. Dieleman, H. Zen et al., "Wavenet: a generative model for raw audio," 2016, <https://arxiv.org/abs/1609.03499>.
- [47] A. Borovykh, S. Bohte, and C. W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," 2017, <https://arxiv.org/abs/1703.04691>.
- [48] Z. Li, J. Han, and Y. Song, "On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning," *Journal of Forecasting*, vol. 39, no. 7, pp. 1081–1097, 2020.
- [49] Z. Pan, Y. Liang, W. Wang, Y. Yu, Y. Zheng, and J. Zhang, "Urban traffic prediction from spatio-temporal data using deep meta learning," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1720–1730, Anchorage, AK, USA, June 2019.
- [50] P. Filonov, A. Lavrentyev, and A. Vorontsov, "Multivariate industrial time series with cyber-attack simulation: fault detection using an lstm-based predictive data model," 2016, <https://arxiv.org/abs/1612.06676>.
- [51] A. Zhao, L. Qi, J. Li, J. Dong, and H. Yu, "LSTM for diagnosis of neurodegenerative diseases using gait data," in *Proceedings of the Ninth International Conference on Graphic and Image Processing (ICGIP 2017)*, vol. 10615, p. 106155B, April 2018.
- [52] X. Jia, A. Khandelwal, G. Nayak et al., "Incremental dual-memory lstm in land cover prediction," in *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 867–876, NS, Halifax, Canada, August 2017.
- [53] C. Huang, C. Zhang, J. Zhao, X. Wu, D. Yin, and N. Chawla, "Mist: A Multiview and Multimodal Spatial-Temporal Learning Framework for Citywide Abnormal Event Forecasting," in *Proceedings of the World Wide Web Conference*, pp. 717–728, CA, San Francisco, USA, May 2019.
- [54] R. Fu, Z. Zhang, and L. Li, "Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction," in *Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation*, pp. 324–328, IEEE, Wuhan, China, November 2016.
- [55] D. Zhang, G. Lindholm, and H. Ratnaweera, "Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring," *Journal of Hydrology*, vol. 556, pp. 409–418, 2018.
- [56] X. Wu, B. Shi, Y. Dong, C. Huang, L. Faust, and N. V. Chawla, "Restful: resolution-aware forecasting of behavioral time series data," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pp. 1073–1082, Torino, Italy, October 2018.
- [57] 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 Advances in Neural Information Processing Systems*, vol. 28, Montreal, Canada, December 2015.
- [58] J.-Y. Wu, M. Wu, Z. Chen, X.-L. Li, and R. Yan, "Degradation-aware remaining useful life prediction with LSTM autoencoder," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.
- [59] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2014.
- [60] Q. Yang, Y. Zhou, Y. Yu, J. Yuan, X. Xing, and S. Du, "Multi-step-ahead host load prediction using autoencoder and echo state networks in cloud computing," *The Journal of Supercomputing*, vol. 71, no. 8, pp. 3037–3053, 2015.
- [61] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep Learning for Solar Power Forecasting-An Approach Using AutoEncoder and LSTM Neural Networks," in *Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, October 2016.
- [62] B. Prenkaj, D. Distanto, S. Faralli, and P. Velardi, "Hidden space deep sequential risk prediction on student trajectories," *Future Generation Computer Systems*, vol. 125, pp. 532–543, 2021.
- [63] Z. Shen, Y. Zhang, J. Lu, J. Xu, and G. Xiao, "A novel time series forecasting model with deep learning," *Neurocomputing*, vol. 396, pp. 302–313, 2020.
- [64] H. Abbasimehr and R. Paki, "Improving time series forecasting using LSTM and attention models," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 1, pp. 673–691, 2022.
- [65] W. Song, Y. Zhang, and S. C. Park, "A novel deep auto-encoder considering energy and label constraints for categorization," *Expert Systems with Applications*, vol. 176, Article ID 114936, 2021.
- [66] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, 2021.
- [67] S. Chaudhari, V. Mithal, G. Polatkan, and R. Ramanath, "An attentive survey of attention models," *ACM Transactions on Intelligent Systems and Technology*, vol. 12, no. 5, pp. 1–32, 2021.
- [68] S. Woo, J. Park, J. Lee, and I. S. Kweon, "Cbam: convolutional block attention module," in *Proceedings of the Computer Vision - ECCV 2018*, pp. 3–19, Munich, Germany, September 2018.
- [69] W. Chen, H. Xu, Z. Chen, and M. Jiang, "A novel method for time series prediction based on error decomposition and nonlinear combination of forecasters," *Neurocomputing*, vol. 426, pp. 85–103, 2021.

Research Article

Algorithm of CAD Surface Generation for Complex Pipe Model in Industry 4.0 Background

Xiaolei Cheng ^{1,2}

¹Intelligent Information Department, Wanbo Institute of Science and Technology, Hefei 230031, Anhui, China

²Hefei University of Technology, Hefei 230009, Anhui, China

Correspondence should be addressed to Xiaolei Cheng; chengxiaolei@wbc.edu.cn

Received 23 February 2022; Revised 8 March 2022; Accepted 11 March 2022; Published 12 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiaolei Cheng. 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 current pipeline surface generation algorithm cannot get the angle information of the corner of complex pipeline surface, which leads to the poor accuracy of the algorithm, the slow speed of 3D point cloud intelligent mosaic, and the large number of effective points. Therefore, a CAD surface generation algorithm for complex pipeline model under the background of Industry 4.0 is designed, extracting and rendering the wireframe model and extracting background of the complex pipeline video. We obtain the angle information of the corner points of the complex pipeline surface, extract and match the feature of the dense point cloud, and construct the 3D point cloud data mosaic model. The pipe surface is generated by using double-nodal B-spline. The experimental results show that the precision and stability of the proposed method are high. In the early stage, the proposed method uses ISS feature extraction algorithm to extract feature of point cloud data, which improves the positioning accuracy effectively and enhances the 3D point cloud intelligent stitching speed.

1. Introduction

Industry 4.0 refers to the Industrial Revolution supported by the Internet and big data, which is the fourth Industrial Revolution derived from the information age. According to the requirements of Industry 4.0, the construction of the factory needs to move forward in the direction of intelligence, establish digital-physical system, make it participate in the purchase of materials, production, and marketing of products and other links, link all links into a whole, comprehensively improve the industrial production efficiency and the degree of product innovation, and then make the Industry get a greater level of progress on the original basis [1, 2]. The strategic goal of Industry 4.0 was put forward by Germany for the first time. At present, China has also realized its importance and has further regulated the pace and direction of industrial development by taking it as a specific standard, so as to indicate a new road for the development of Chinese Industry. With the development of society and economy and the improvement of science and technology, the Internet has

become the focus of development and application in various fields.

Tubular engineering structure has the advantages of good mechanical properties and light weight. It is widely used in construction, vehicle engineering, biomedicine, chemical Industry, national defense and military, and other fields, for example, frame structures in buildings, lightweight engineering in the field of vehicles, vascular simulation in medicine, deep-sea drilling, pipeline robots, and other major equipment [3–5]. In the field of computer-aided engineering (CAE), in order to gain a better understanding of the physical properties of objects, it is often necessary to simulate precisely the objects studied. Accurate simulation can shorten the development cycle, reduce costs, and reduce risks, and accurate modeling is the basis of accurate simulation [6, 7]. Most tubular structures are shell structures. Their accurate simulation has important theoretical and practical value.

Reference [8] proposes an automatic DSM generation method based on CAD model. Firstly, the mates feature information of the top-level components in the structural

feature tree of CAD model in SolidWorks platform is extracted by SolidWorks API, arranged in a specific order, and stored in the database. Secondly, the geometric relationship between assembled parts is analyzed to determine the influence of various types of fit on the connection relationship between parts, and the corresponding analysis and comparison rules are established. Finally, the design of automatic generation method of pipe surface is successfully completed by using Visual Basic. Reference [9] proposed a surface equivalent segmentation method with user-defined conditions as the core. The user designs a set of conditions in advance according to the segmentation problem and related surface point attributes, calculates the boundary points based on this, searches the boundary points in turn along the growth direction of the boundary edge, and then generates the boundary edge by ordered boundary point spline interpolation. By constructing the boundary edges of the surface and its subpatches, the surface is divided into a group of isosurfaces. In order to ensure the generality of segmentation, an accurate calculation method of initial boundary points and boundary points of sampling units in parameter domain is proposed. Reference [10] proposes a method to obtain point cloud data based on secondary development of UG NX platform. This paper introduces the tools and basic characteristics of UG NX secondary development and comprehensively uses the secondary development tools and Visual Studio 2010 development environment in UG NX10 0 is designed to complete the dialog box of CAD model discretization. This paper expounds the principle and basic flow of CAD model discretization, compiles the callback function in the dialog box, realizes the discretization function, and obtains the point cloud data. Reference [11] designed an oil and gas pipeline data monitoring and analysis model based on big data management architecture. Firstly, they build the overall architecture of the model, collect and sort out the pipeline data through big data support, preprocess it, and determine the number of rows and columns in the associated arrangement. Brillouin sensing technology is introduced. Based on distributed optical fiber sensors and sensing data, optical fiber is used to replace the data composed of hundreds of traditional sensing points to generate feedback sensing sequence. The distance data is generated by calculating the template and optical fiber feedback, and the distance values of the current oil and gas pipeline are determined by relying on the scattering point emission spectrum relationship distance. Through comparison, whether the current pipeline has bending can be determined to realize the monitoring of pipeline data.

However, the above methods can not obtain the included angle information of corner points of complex pipeline surface, resulting in poor accuracy of the algorithm, slow speed of 3D point cloud intelligent splicing, and more effective points of surface splicing. Therefore, the CAD surface generation algorithm of complex pipeline model under the background of Industry 4.0 is designed. The wireframe model of CAD surface of complex pipeline model is extracted and drawn, and the complex pipeline video is collected according to the model to extract the video

background. Then the dense point cloud features of pipe surface are extracted and matched, and the 3D point cloud data stitching model is constructed. Finally, the pipe surface is generated by multiple node B-spline. Finally, experiments show that the algorithm has higher accuracy and stability and lower error rate and can meet the needs of CAD surface generation of complex pipeline model.

2. CAD Surface Generation Algorithm of the Complex Pipeline Model

2.1. Extracting and Drawing the Wireframe Model. Line frame model of geometric modeling system is the earliest technology used to describe object model, which plays an important role in the early development stage of products [12]. It not only can reflect the designer's innovative thinking, but also has the advantages of flexible operation. For example, in the conceptual design of automobiles, in order to shorten the R&D cycle, designers usually draw sketches to represent the shape of the car. In the early stage of underground pipeline design, designers often use wireframe model to replace tubular structure model [13]. The mainstream software has wireframe processing function, which simplifies the tubular model into wireframe model through curve and surface operations such as average, offset, and mixing. Figure 1 shows the pipe wireframe.

The extracted wireframe structure is divided into multiple two joints and three joints, as shown in Figures 2 and 3. According to the curve bending degree and actual accuracy requirements, determine the number of segments, as shown in Figure 2, so that the number of curves in each joint is 2, as shown in Figure 3, so that the number of curves in each joint is 3 and the node vector is $U = \{0, 0, 0, 0.5, 1, 1, 1_n\}$, and store the geometric information of each joint in the form of text file.

2.2. Complex Pipeline Video Acquisition. TKF240 is selected as the acquisition module of the surveillance camera to complete the functions of the control exposure, image gain, and white balance of the complex pipeline video [14, 15]. The functional block diagram of the TKF240 acquisition sensor is shown in Figure 4.

As can be seen from Figure 4, the photosensitive array of TKF240 acquisition sensor includes 1632×1232 pixels, up to 2 million pixels. The video signal processing module of the acquisition sensor is the core of the whole module. The acquisition sensor is used to sharpen the edge of the complex pipeline video image.

In order to ensure the integrity of the video data of complex pipelines, it is necessary to use a video collector for complex pipelines to store the video data of complex pipelines. The 32-bit processor shall be used in the video monitoring system for complex pipelines oriented to Industry 4.0, and the video collector for complex pipelines shall also choose 32-bit processors. Considering that the video collector is larger than 8M to ensure the integrity of the video data of complex pipelines, the SCD985 chip is selected, and the single chip has a capacity of 16×16 . Since the video

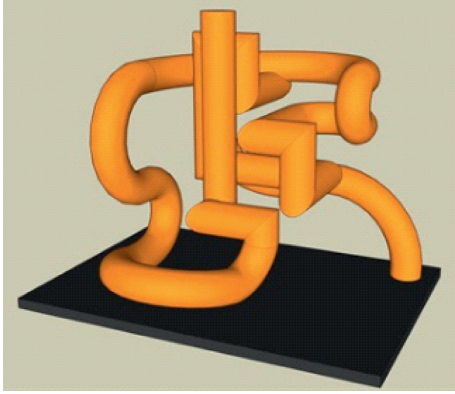


FIGURE 1: Wireframe model.



FIGURE 2: Two-joint pipe.

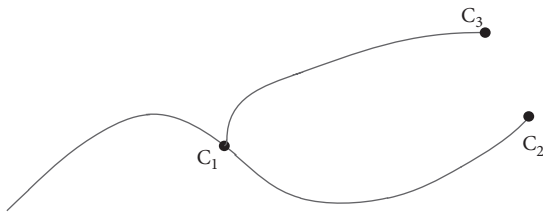


FIGURE 3: Three-joint pipe.

data transmission of 32-bit is required in the motion video monitoring system, the 16- \times 32 collector composed of two pieces shall be used.

If the video image of complex pipeline captured by the monitor camera is displayed in the monitor terminal, it must be realized through the transmission interface of the monitor network. The Ethernet of Industry 4.0 is used to transmit the video image of complex pipeline. The monitoring terminal is used as the server and the video collector of complex pipeline as the client. Monitoring network transmission interface is the data interface connecting complex pipeline video controller and media access controller and can support the independent interface of complex pipeline video [16]. In the system register, the complex pipeline video data transmission is realized by configuring the information technology and computer technology of Industry 4.0. The monitoring network transmission interface is illustrated in Figure 5.

The Monitor Network Transmission Interface allows monitoring applications to access registers. The Monitor Network Transmission Interface based on Industry 4.0 supports multiple input modes and supports complex pipe video images with a maximum input resolution of 8192 \times 8192, monitoring network transmission interface with video data scaling and previewing video images and other functions, but also supporting the rotation and flipping functions. Complex piping video surveillance systems can

capture complex piping video through the transmission interface of the surveillance network, and surveillance cameras with industrial 4.0 designs have super-high pixels of 3 million to meet the requirements of motion video resolution [17, 18].

Based on the advantages of the Fourth Industrial Revolution, the paper integrates several video processing units in the hardware platform of the monitoring system, collects the image data of the complex pipeline video on the collector sensor, and completes the design of the video collector of the complex pipeline and develops the monitoring network transmission interface of the monitoring system by using Ethernet, completes the design of the monitoring network transmission interface by analyzing the performance and circuit interface of the monitoring network, and realizes the hardware design of the moving video monitoring system [19, 20].

2.3. Extracting the Complex Pipeline Video Background.

In the context of Industry 4.0, to extract the background of complex pipeline video, we must first determine the picture pixels in the complex pipeline video background. In the complex pipeline video collected by the collector, the value of each pixel follows a specific law. For complex pipeline video monitoring system, due to different monitoring scenes, the background of complex pipeline video usually presents single-mode characteristics [21, 22].

Suppose that, for a complex pipeline video frame sequence $Q = \{q_1, q_2, \dots, q_n\}$, the pixel values in the complex pipeline video together constitute a pixel set $W = \{w_1, w_2, \dots, w_n\}$, so that μ is the mean value of pixel set W and δ is the variance of pixel set W ; when the complex pipeline video pixels in W conform to the normal distribution, there is

$$W = \frac{1}{\sqrt{2\pi}\delta} e^{-(p-\mu)^2/2\delta^2}. \quad (1)$$

When extracting the complex pipeline video background, the surveillance camera is usually static, and the extracted complex pipeline video background value is the mean value of the statistical value. Assuming that the extracted complex pipeline video background is $\aleph(x, y)$ and the complex pipeline video frames collected by the system are $f_i(x, y)$ and ($i = 1, 2, 3, \dots, N$), the extracted complex pipeline video background is

$$\aleph(x, y) = \frac{\sum_{i=1}^N f_i(x, y)}{N}. \quad (2)$$

In a complex pipeline video surveillance system, when the surveillance camera oscillates slightly, the captured complex pipeline video pixels will not follow the normal distribution, and the above method will cause the extracted complex pipeline video background confusion [23]. Based on the normal distribution, a coefficient α representing the speed of pixel update is introduced to control the proportion of complex pipeline video background. The updated complex pipeline video background extraction formula is as follows:

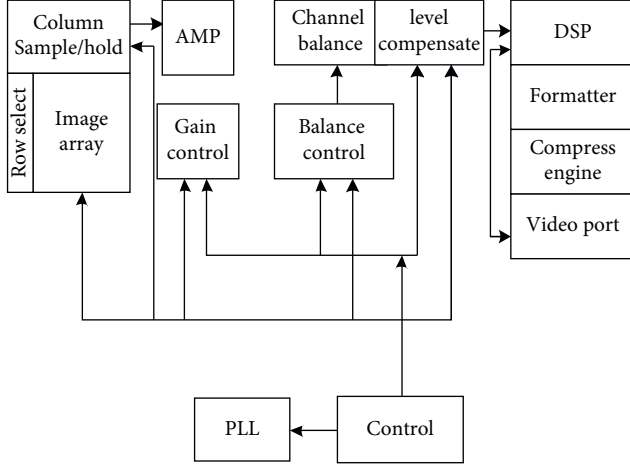


FIGURE 4: Functional block diagram of the TKF240 acquisition sensor.

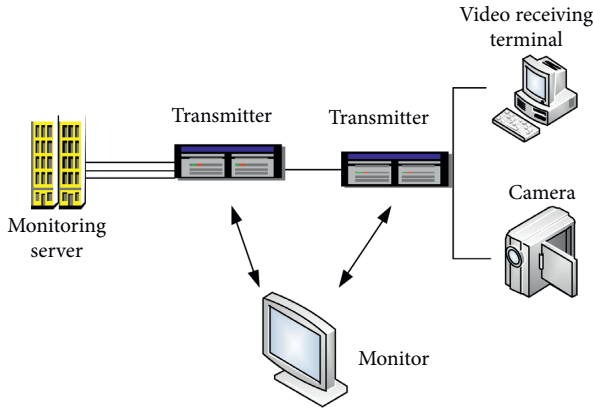


FIGURE 5: Principle of the Monitor Network Transmission Interface.

$$\aleph(x, y) = \alpha[\mu_{i-1} \cdot f_i(x, y)]. \quad (3)$$

Because it is difficult for the original complex pipeline video background extraction formula to obey the normal distribution, the complex pipeline video background extraction formula is updated by introducing the coefficient of update speed.

2.4. Feature Extraction and Matching of Dense Point Cloud on the Pipe Surface. In the process of reconstruction of complex pipeline surface, the included angle information of corner points of complex pipeline surface is firstly obtained by combining SUSAN corner detection theory, and corner features under contrast of dense point clouds of different pipeline surfaces are extracted [24, 25]. Surface initial matching was carried out on extracted corner point features to obtain the initial matching point set, and the matching results were optimized with the least square median theory. The specific process is described as follows:

In the complex pipe surface template area, the maximum and nonzero point is selected as the corner point, where g represents a fixed threshold value in the dense point cloud of

the surface, and \aleph represents the minimum contrast of corner points that can detect the surface point cloud.

The gray value of complex pipeline surface was analyzed, and the adaptive threshold under contrast \aleph of point cloud of different surfaces was extracted, which was expressed by

$$\aleph = a \cdot \frac{\sum_{i=1}^n I_{i\max} - \sum_{i=1}^n I_{i\min}}{n}, \quad (4)$$

where $I_{i\max}$ and $I_{i\min}$, respectively, represent the maximum i gray values and the minimum i gray values in the point cloud of pipe surface.

Firstly, the initial matching point (m_{1i}, m_{2j}) of the point cloud of the pipe surface is obtained by the correlation method. The two corners are centered on m_{1i} and m_{2j} , respectively, and R is the neighborhood $N(m_{1i})$ and $N(m_{2j})$ of the radius of the pipe surface. Pipe surface point cloud matching points m_1 and m_2 are associated points, and pipe surface is 3×3 . If F in the matrix is the module matrix of two pipe surface point clouds, and the surface point cloud module matrix vector \tilde{m}_i is the homogeneous coordinate of the matching point m_i , there are

$$\tilde{m}_i^T F \tilde{m}_1 = 0. \quad (5)$$

Assuming that the proportion of wrong matching in the whole pipeline surface point cloud matching set is ε , the probability that each pipeline surface point cloud sample contains U pairs of correct matching points in multiple pipeline surface point cloud samples is given by

$$U = 1 - (1 - \varepsilon)^{p^m}. \quad (6)$$

2.5. 3D Point Cloud Data Mosaic Model. Set the quadratic pipe surface equation composed of the i feature point (x, y) , and the general form is

$$z_i = \begin{cases} f(x, y), \\ a_0 + a_1x + a_2y + a_3x^2 + a_4y^2, \end{cases} \quad (7)$$

where a_0, a_1, a_2, a_3, a_4 represent the rotation angle. The minimum value is obtained according to the least square principle, namely,

$$\varepsilon^2 = \sum_i (a_0 + a_1x + a_2y + a_3x^2 + a_4y^2 - z_i)^2. \quad (8)$$

Derive formula (8) so that its final value is 0. Then, it can be combined with formula (8) to obtain the fitting equation of quadratic pipe surface. The specific expression is as follows:

$$\chi(x, y) = (x, y, a_0 + a_1x + a_2y + a_3x^2 + a_4y^2). \quad (9)$$

Point cloud mosaic mainly refers to the use of arbitrary transformation to align the point cloud data in two groups of different regions, so that point clouds obtained from points in each region can be accurately matched and spliced [26–28]. Several common point cloud stitching strategies are given as follows:

- (1) Sequence splicing

(2) Global stitching

Three-dimensional topography measurement is mainly composed of control computers, industrial robots, and topography [29, 30]. Among them, the point cloud mosaic model based on the world coordinate system is mainly composed of the following coordinate systems:

- (1) The coordinate system of the topography sensor is $O_s - X_s Y_s Z_s$
- (2) Frame coordinate system composed of rigid connection of topography sensor is $O_f - X_f Y_f Z_f$
- (3) The world coordinate system is $O_w - X_w Y_w Z_w$

Any two coordinate systems between the above three coordinates can be converted, P stands for any point within the set range, and the coordinate of this point in the topography sensor coordinate is P_s . Through transformation relation ${}^f_s A$, point cloud measurement data can be obtained and converted to the world coordinate system, and the corresponding relation can be obtained as follows:

$$P_w = {}^w_s A \cdot {}^f_s A \cdot P_s. \quad (10)$$

From the analysis formula, it can be seen that the point cloud data obtained by the robot driving the profilometric sensor at different measuring stations is converted to the world coordinate system, which is constructed on the laser transmitter [31]. Because the transmitter needs to keep the position unchanged in the course of measurement, it is necessary to make sure that the robot can work out the trajectory in advance and measure different positions according to the trajectory.

The swarm intelligence algorithm was used to solve the 3D point cloud data mosaic model [32]: Feature points were extracted for the initial point cloud s_1 and target point cloud s_2 to obtain the simplified model. The swarm intelligence optimization algorithm was used to optimize the evaluation function and obtain the transformation matrix. The obtained transformation matrix is applied to the initial point cloud s_1 to realize the splicing with point cloud s_2 . The specific operation process is shown in Figure 6:

- (1) Set up point sets corresponding to different point clouds, and determine the number of final feature points according to the size of data scale
- (2) Structured processing of point sets through the K-D tree method to accelerate the acquisition of the distance between point clouds [33]
- (3) A K-D tree is set up for all points in the source point set g by using the target point set to quickly obtain K neighborhood points and calculate the distances between different point clouds and K neighborhood points, respectively
- (4) The regional point sets composed of each point and K neighborhood shall be counted, the point expansion matrix shall be established, and the corresponding feature vectors shall be calculated
- (5) The eigenvalue is evaluated through the eigenpoint extraction algorithm, and the corresponding eigenpoints are extracted

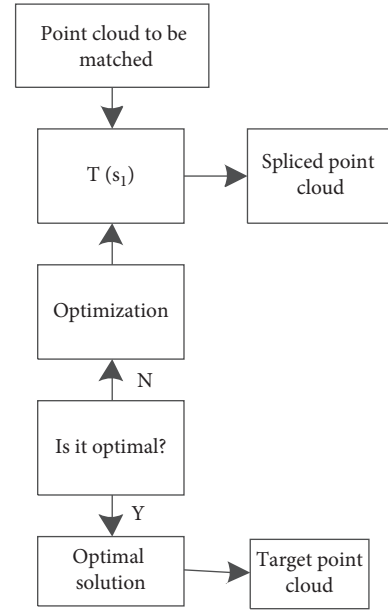


FIGURE 6: Flowchart of point cloud data splicing.

- (6) The simplified model after extracting the feature points is regarded as a new initial model, and the swarm intelligence algorithm is used to determine the initial parameters and convergence conditions
- (7) Initialize the optimized individuals to ensure the corresponding spatial location of each search field
- (8) When an iteration is completed, the current optimal solution and the global optimal solution shall be compared, the better optimal solution shall be set to all new optimal solutions, and step (7) shall be returned, and the above operation process shall be repeated until a required number of iterations are completed
- (9) Comparing each variable of the final optimal solution with six parameters of the transformation matrix, the transformation matrix is applied to G , and obtaining a new point cloud is the point cloud after the completion of the mosaic

2.6. Pipeline Surface Generation Based on Double-Nodal B-Spline. In the process of generating pipeline surface, the free deformation technology is used to place the deformed part of the 3D model of pipeline surface parameters into a specific hexahedron space and assign fixed grid parameter coordinates to all points in the hexahedron and on the hexahedron boundary and promote the model to produce deformation effect by adjusting control points. The parametric optimization generation of pipeline surface is effectively realized [34, 35]. The specific steps are as follows:

In the process of parametric optimization generation of pipe surface, it is assumed that $r_{i,j,k}$ represents all marked pipe surface mesh control vertices, and the part of the pipe surface parameter 3D model to be modeled is placed in a specific hexahedron space based on $L_{i,2}$ and $L_{i,N+1}$; then all

marked pipe surface mesh control vertices shall meet the following conditions:

$$r_{i,j,k} = U + V \cdot jL_{i,2} + L_{i,N+1} \cdot W. \quad (11)$$

Here, i , N , and j represent the boundary of pipe surface mesh control vertices, and U , V , and W represent the coordinates of pipe surface mesh control vertices [36].

In the process of generating pipe surface, the model is deformed by adjusting the control points, and the new coordinates of point Q after deformation are calculated by using the following formula [37]:

$$\gamma(u, v, w) = \frac{b_{i,j}(u) \cdot b_{j,m}(v) b_{k,n}(w)}{r_{i,j,k} \cdot v}. \quad (12)$$

Here, $r_{i,j,k}$ represents the position of the new coordinates of the control vertex, and $b_{i,j}(u)$, $b_{j,m}(v)$, and $b_{k,n}(w)$ represent the basis function of the pipe surface, respectively, which completes the generation of the pipe surface.

3. Experimental Design and Result Analysis

The application effectiveness of the proposed algorithm is verified by the test object with a complex three-tube structure.

3.1. Experimental Setup. In this experiment, the three mutual components shown in Figure 7 are used. The tubular surface structure is reconstructed from the wireframe model of the corner. Firstly, the geometric information such as control points and weight factors of the three lines is extracted and stored in the form of files. On the cross section perpendicular to the curve, the weight factor of the control point is $W = \{1, 2^{1/2}, 1, 2^{1/2}, 1, 2^{1/2}, 1\}$ and the node vector is $V = \{0, 0, 0, 1, 1, 1\}$. Figure 7 shows the reconstructed three continuous NURBS surfaces and the corresponding control network, with a radius of 0.3 mm.

According to the mutual formation mode of Figure 7, the CAD surface of the interface pipe shown in Figure 8 is generated by computer software. Among them, Table 1 shows the computer configuration and running time during the experiment.

In order to prove the comprehensive effectiveness of the complex pipeline surface reconstruction method based on genetic algorithm, a simulation experiment is needed. In the process of complex pipeline surface reconstruction, firstly, the genetic parameters are given, the initial population size $\text{Num} = 50$, and the maximum number of iterations $\text{MAXGEN} = 60$. Binary coding is used for coding, and the coding accuracy is 0.01. The local pipeline surface to be reconstructed in the first group of reverse engineering is shown in Figure 9.

3.2. Analysis of Matching Effect of Surface Reconstruction. The genetic algorithm is used for optimization to find the optimal position for the reconstruction and matching of complex pipeline surfaces. After the coordinate rotation operation, the coordinates of the obtained complex pipeline

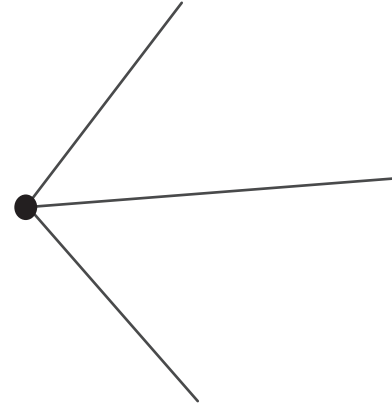


FIGURE 7: Three straight lines forming 60° to each other.

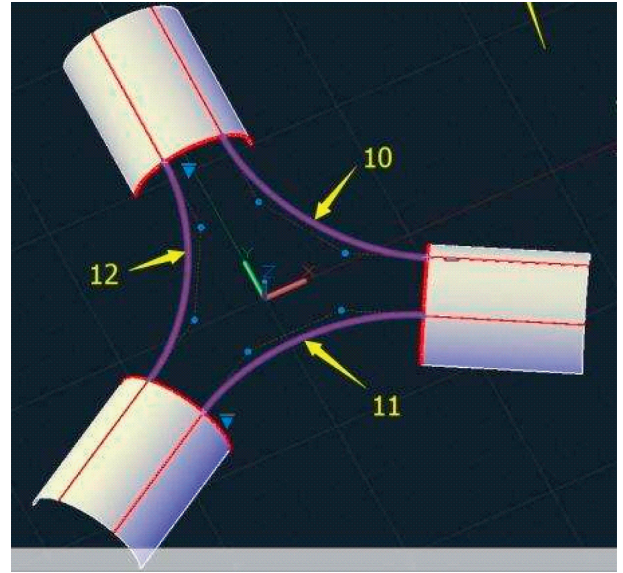


FIGURE 8: Interface pipe CAD surface generation effect.

surfaces are transformed to realize the matching of local complex pipeline surfaces in reverse engineering. The reconstruction and matching effect pictures of local pipeline surfaces when the genetic algorithm is iterated 40 times are given, respectively, as shown in Figure 10.

By analyzing Figure 10, when using genetic algorithm for rough matching of point cloud of complex pipeline surface, with the continuous increase of iteration times, the matching result will eventually converge to near the global optimal solution, which provides a good initial solution for further realizing complex pipeline surface reconstruction. The effect diagram of global complex pipeline surface matching in reverse engineering is shown in Figure 11.

3.3. Performance Test. Comparative experiments are carried out by using the automatic surface generation method based on CAD model proposed in [8], the automatic surface generation method based on surface equivalent segmentation method, and the improved algorithm proposed in [9]. Under different experimental times, the accuracy (%), error

TABLE 1: Computer configuration and running time of experimental environment.

Computer configuration	Number of joints	Number of control points	Running time/s
Computer core Quad main frequency 2.66HZ memory 8GB	15	750	5
	25	1290	10
	35	1550	17
	45	2150	25
	55	2800	28
	65	3500	36

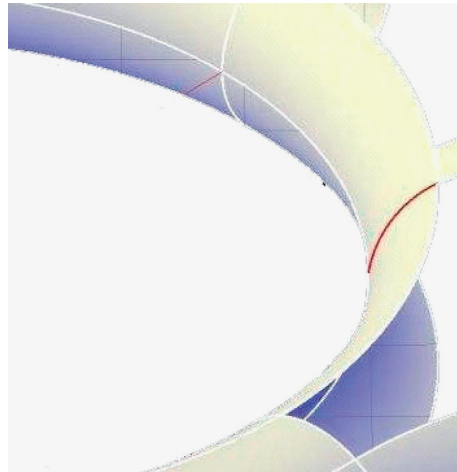


FIGURE 9: Surface of local pipeline to be reconstructed.

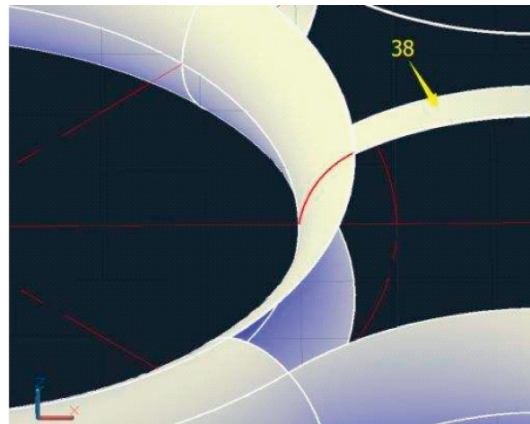


FIGURE 10: Surface of pipeline surface matching effect in 40 iterations.

rate (%), and stability (%) of the surface generation algorithms established by the three methods are compared, respectively. The comparison results are used to measure the overall effectiveness of the three algorithms in establishing the parametric three-dimensional model of pipe surface.

It can be seen from Figures 12 and 13 that the accuracy of the algorithm in this paper is close to 100%, and the error is less than 3%, which is higher than the comparison method. The main reason is that the algorithm in this paper extracts and draws the wireframe model of pipeline

surface, extracts the background of complex pipeline video, optimizes the accuracy of surface generation, and reduces the error of surface generation. As can be seen from Figure 14, the stability of the algorithm in this paper is higher than 97%, and the stability of the comparison algorithm is higher than 85%, but in comparison, the stability of the algorithm in this paper is higher. According to the results of Figure 12 and Figure 14, the overall effectiveness of the pipe surface established by the improved algorithm is better than the traditional

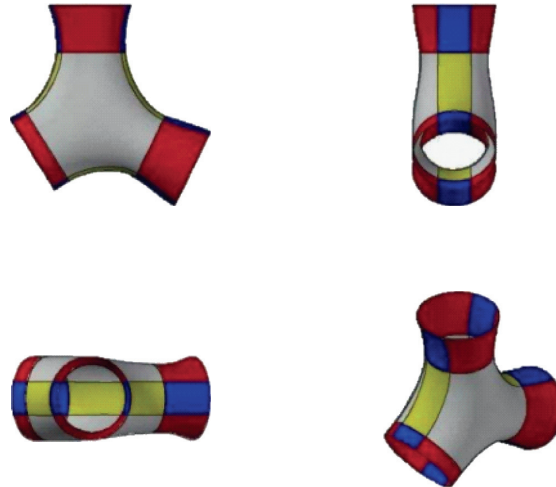


FIGURE 11: Global complex pipeline surface matching effect.

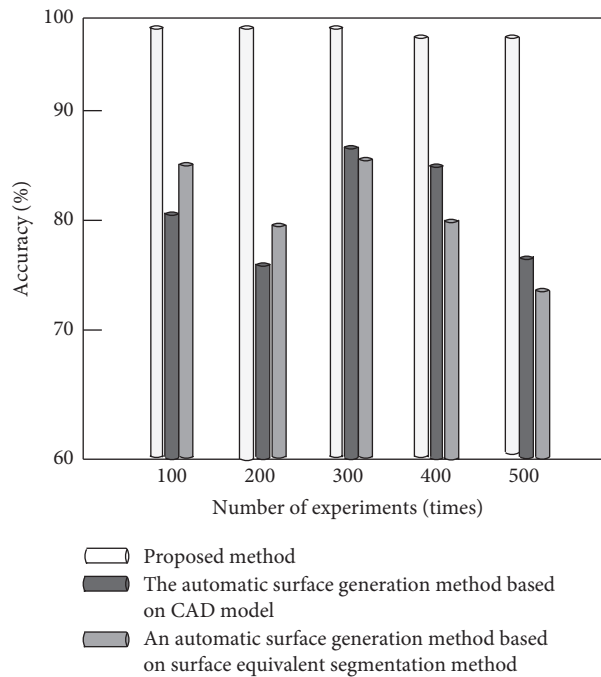


FIGURE 12: Comparison of accuracy of different methods.

algorithm. It is also proved that the improved algorithm can splice smooth and continuous pipe surfaces by using curve and surface theory.

3.4. 3D Point Cloud Intelligent Splicing Speed Test. In order to further verify the advantages of the proposed method, the following experimental tests compare the three-dimensional point cloud intelligent splicing speed of four different methods. The specific experimental comparison results are shown in Figure 15.

By analyzing the experimental data in Figure 15, it can be seen that the proposed method not only effectively improves the positioning accuracy, but also comprehensively enhances the three-dimensional point cloud

intelligent splicing speed of the whole method through the ISS feature extraction algorithm in the early stage. Therefore, the splicing speed of the algorithm in this paper is less than 8, and the maximum splicing speed of the comparison method is 16 and 13. Therefore, the 3D point cloud intelligent splicing speed of this method is obviously better than the other three methods.

3.5. Test of Effective Splicing Points. The number of effective stitching points is also an important index to verify the comprehensive performance of each method. The higher the number of effective stitching points is, the more ideal stitching results will be obtained. Table 2

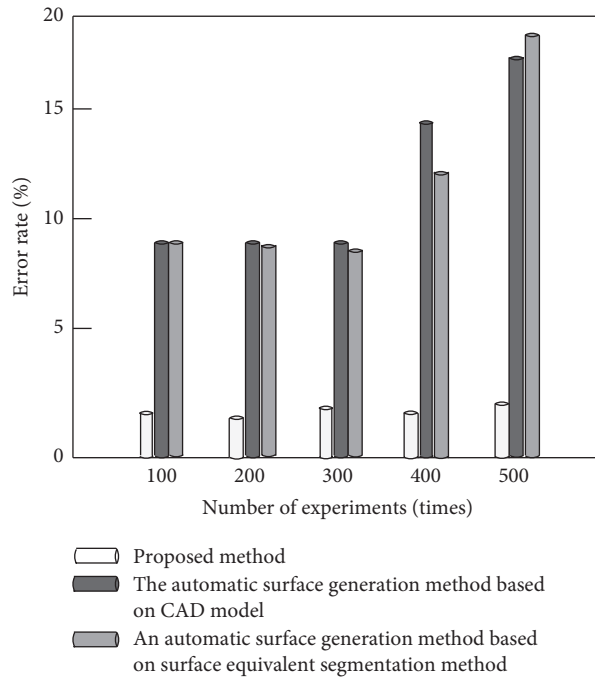


FIGURE 13: Comparison of error rates of different methods.

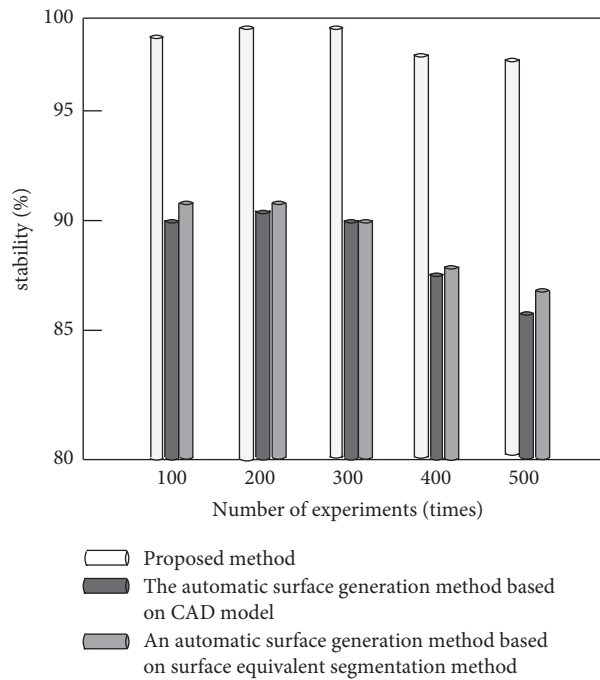


FIGURE 14: Stability comparison of different methods.

shows the comparison results of the effective points of three different methods.

By analyzing the experimental data in Table 2, it can be seen that, with the continuous increase of rotation angle, the number of stitching effective points of each method is decreasing, but the number of stitching

effective points of the proposed method is significantly higher, with a minimum of 137, which is higher than 98 in [8] and 83 in [9], mainly because the proposed method accurately extracts the feature points of point cloud data, effectively promoting the improvement of comprehensive performance of the method.

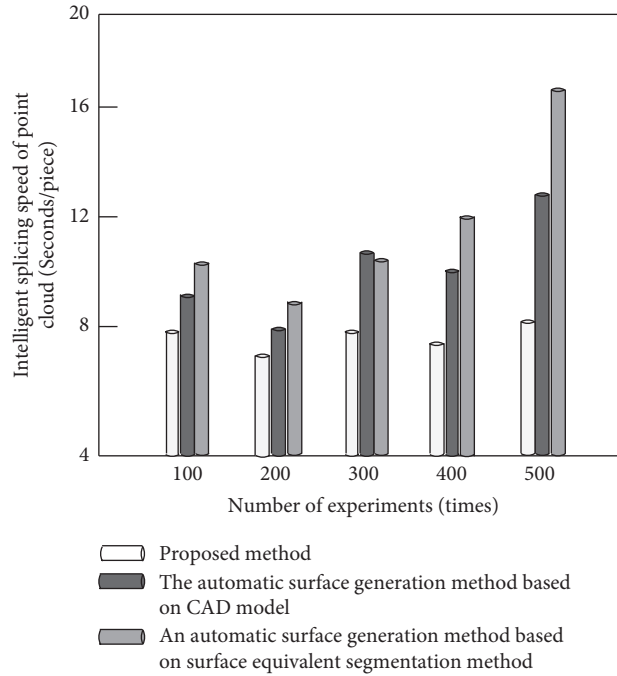


FIGURE 15: Comparison results of 3D point cloud intelligent splicing speed of different methods.

TABLE 2: Comparison results of effective splicing points of different methods.

Rotation angle/°	Number of spliced effective points		
	Proposed method	Reference [8] method	Reference [9] method
15	145	121	113
20	140	116	105
25	137	111	97
30	134	107	92
35	130	102	88
40	137	98	83

4. Conclusion

In order to optimize the surface reconstruction effect of complex pipeline, the CAD surface generation algorithm of complex pipeline model under the background of Industry 4.0 is designed. On the basis of extracting and matching the characteristics of dense point cloud of pipeline surface under different contrast of dense point cloud of pipeline surface, a three-dimensional point cloud data stitching model is constructed, and the pipeline surface generation algorithm is designed based on multiple node B-spline. The experimental results show that the proposed method has high accuracy and stability, effectively improves the positioning accuracy of surface point cloud data, and comprehensively enhances the intelligent splicing speed of 3D point cloud of the whole method.

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 no conflicts of interest regarding this work.

Acknowledgments

This work was supported by Key Project of Natural Science Research in Colleges of Anhui Provincial Department of Education, “Research and Implementation of Mobile Teaching Management Platform under J2EE Framework” (KJ2020A1170) and the research was supported by Key Project of Natural Science Research in Colleges of Anhui Provincial Department of Education “Research on gray uneven image segmentation algorithm based on Cauchy distribution” (KJ2020A1171).

References

- [1] V. S. Osipov, E. I. Larionova, and A. V. Sigarev, “The key types of economic activities in the pleasure economy in the context of developing countries’ transition to Industry 4.0,” *International Journal of Trade and Global Markets*, vol. 14, no. 4/5, p. 413, 2021.

- [2] S. Mosser, "Maintenance modelling and optimization in the industry 4.0 context: applications to the maintenance and design of industrial vehicles[J]," *Sustainable Engineering and Innovation* ISSN, vol. 3, no. 2, pp. 130–138, 2021.
- [3] S. Arslan, M. Eyvaz, S. Güllü, A. Yükksekda, Koyuncu, and E. Yüksel, "Pressure assisted application of tubular nanofiber forward osmosis membrane in membrane bioreactor coupled with reverse osmosis system," *Journal of Water Chemistry and Technology*, vol. 43, no. 1, pp. 68–76, 2021.
- [4] N. T. Vinh and A. T. Le, "Simulating behavior and evaluating rational parameters of the tubular steel joint using flanges and bolts by FEM analysis," *Science & Technology Development Journal - Engineering and Technology*, First, vol. 2, no. 4, , 2020.
- [5] C. B. Nayak, "Experimental and numerical investigation on compressive and flexural behavior of structural steel tubular beams strengthened with AFRP composites - ScienceDirect," *Journal of King Saud University - Engineering Sciences*, vol. 33, no. 2, pp. 88–94, 2020.
- [6] J. Shi, H. Zhang, Q. Wang, Z. Duan, and Z. Chen, "Biomimetic rigid cryogels with aligned micro-sized tubular structures prepared by conventional redox-induced cryopolymerization in a freezer," *Chemical Engineering Journal*, vol. 427, Article ID 131903, 2021.
- [7] D. N. Serras, S. D. Panagaki, K. A. Skalomenos, and G. D. Hatzigeorgiou, "Inelastic lateral and seismic behaviour of concrete-filled steel tubular pile foundations," *Soil Dynamics and Earthquake Engineering*, vol. 143, no. 1, Article ID 106657, 2021.
- [8] Z. K. Li and W. W. Yin, "An automatic design structure matrix generation system based on 3D CAD model," *Computer Engineering and Science*, vol. 42, no. 3, pp. 483–492, 2020.
- [9] J. C. Hu and G. L. Zheng, "CAD surface iso-segmentation," *Journal of Computer-Aided Design & Computer Graphics*, vol. 33, no. 1, pp. 153–160, 2021.
- [10] H. Y. Xiang, Z. R. Zhou, L. P. Ma, and C. J. Leng, "Application of UG NX secondary development on CAD model discretization," *Machine Tool & Hydraulics*, vol. 48, no. 05, pp. 151–154+193, 2020.
- [11] X. Sun, S. Q. Li, R. Xie, J. Q. Li, and J. S. Hao, "Oil and gas pipeline data monitoring and analysis model based on big data management framework," *Modern Electronics Technique*, vol. 43, no. 17, pp. 102–105, 2020.
- [12] Y. Tian, Q. Qian, Y. Sheng et al., "High energy density in poly(vinylidene fluoride-trifluoroethylene) composite incorporated with modified halloysite nanotubular architecture," *Colloids and Surfaces A Physicochemical and Engineering Aspects*, vol. 625, Article ID 126993, 2021.
- [13] P. Li, Y. Yu, and D. Xu, "Research on optimization design of underground garage comprehensive pipeline based on BIM technology[J]," *IOP Conference Series: Materials Science and Engineering*, vol. 780, no. 5, Article ID 052008, 2020.
- [14] J. Pérez-Bailón, B. Calvo, and N. Medrano, "1.0 V-0.18 m CMOS tunable low pass filters with 73 dB DR for on-chip sensing acquisition systems[J]," *Electronics*, vol. 10, no. 5, p. 563, 2021.
- [15] A. Lay-Ekuakille, J. D. Okitadiowo, M. A. Ugwiri, S. Maggi, and G. Passarella, "Video-sensing characterization for hydrodynamic features: particle tracking-based algorithm supported by a machine learning approach," *Sensors*, vol. 21, no. 12, p. 4197, 2021.
- [16] P. Y. Kong, J. C. Wang, K. S. Tseng, Y.-C. Yang, Y.-C. Wang, and J.-A. Jiang, "An adaptive packets hopping mechanism for transmission line monitoring systems with a long chain topologyVideo-Sensing Characterization for Hydrodynamic Features: particle Tracking-Based Algorithm Supported by a Machine Learning Approach," *International Journal of Electrical Power & Energy Systems*, vol. 124, Article ID 106394, 2021.
- [17] P. F. Lamas, M. C. Echarri, and F. F. Lanas, "Design and empirical validation of a Bluetooth 5 fog computing based industrial CPS architecture for intelligent industry 4.0 shipyard workshops," *IEEE Access*, vol. 8, pp. 45496–45511, 2020.
- [18] L. Gogolák and I. Fürstner, "Wireless sensor network aided assembly line monitoring according to expectations of industry 4.0[J]," *Applied Sciences*, vol. 11, no. 1, p. 25, 2020.
- [19] A. D. Neal, R. G. Sharpe, K. V. Lopik, J. Tribe, and A. A. West, "The potential of industry 4.0 Cyber Physical System to improve quality assurance: an automotive case study for wash monitoring of returnable transit items," *CIRP Journal of Manufacturing Science and Technology*, vol. 32, pp. 461–475, 2021.
- [20] S. Cavalieri and M. G. Salafia, "Insights into mapping solutions based on OPC UA information model applied to the industry 4.0 asset administration shell[J]," *Computers*, vol. 9, no. 2, p. 28, 2020.
- [21] M. Pieroni and E. Barausse, "Foreground cleaning and template-free stochastic background extraction for LISA," *Journal of Cosmology and Astroparticle Physics*, vol. 2020, no. 7, p. 021, 2020.
- [22] J. Zuo, Z. Jia, and J. Yang, "Moving object detection in video sequence images based on an improved visual background extraction algorithm," *Multimedia Tools and Applications*, vol. 79, no. 39, 40 pages, 2020.
- [23] M. A. Yasir and Y. H. Ali, "Review on real time background extraction: models, applications, environments, challenges and evaluation approaches," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 17, no. 2, p. 37, 2021.
- [24] H. Yoo, A. Choi, and J. H. Mun, "Acquisition of point cloud in CT image space to improve accuracy of surface registration: application to neurosurgical navigation system," *Journal of Mechanical Science and Technology*, vol. 34, no. 6, pp. 2667–2677, 2020.
- [25] S. Boljanovi and A. Carpinteri, "Computational analysis of a surface corner crack under cyclic loading," *Procedia Structural Integrity*, vol. 28, no. 1, pp. 2370–2377, 2020.
- [26] J. Gao, F. Li, C. Zhang, W. He, J. He, and X. Chen, "A method of D-type weld seam extraction based on point clouds[J]," *IEEE Access*, vol. 9, pp. 65401–65410, 2021.
- [27] F. Lu, N. Kothari, X. Feng, and L. Zhang, "Equivalence classes in matching covered graphs[J]," *Discrete Mathematics*, vol. 343, no. 8, Article ID 111945, 2020.
- [28] B. Anne, G. Louis, K. Matthias, V. O. Goddert, and M. Hans-Gerd, "Automatic extraction and measurement of individual trees from mobile laser scanning point clouds of forests," *Annals of Botany*, vol. 128, no. 6, pp. 787–804, 2021.
- [29] B. Cai, L. Liu, J. Wu, X. Chen, and Y. Wang, "Three-dimensional shape measurement based on spatial-temporal binary-coding method," *Measurement Science and Technology*, vol. 32, no. 9, Article ID 095014, 2021.
- [30] Y. Arai, "Precise wide-range three-dimensional shape measurement method to measure superfine structures based on speckle interferometry," *Optical Engineering*, vol. 59, no. 1, p. 1, 2020.
- [31] W. Sun, J. Wang, and F. Jin, "An automatic coordinate unification method of multitemporal point clouds based on virtual reference datum detection," *Ieee Journal of Selected*

- Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3942–3950, 2020.
- [32] R. S. Rampriya and R. Suganya, “Segmentation of 3D point cloud data based on supervoxel technique,” *Procedia Computer Science*, vol. 171, pp. 427–435, 2020.
- [33] W. R. Min, M. O. Sang, J. K. Min, H. H. Cho, and T. H. Kim, “Algorithm for generating 3D geometric representation based on indoor point cloud data,” *Applied Sciences*, vol. 10, no. 22, p. 8073, 2020.
- [34] D. V. Lozhkin, P. V. Maksimov, K. V. Fetisov, and A. A. Krotkikh, “Topological and parametric optimization in the design of SLM products,” *Russian Engineering Research*, vol. 41, no. 9, pp. 851–853, 2021.
- [35] F. Casino, P. Lopez-Iturri, E. Aguirre, L. Azpilicueta, F. Falcone, and A. Solanas, “Enhanced wireless channel estimation through parametric optimization of hybrid ray launching-collaborative filtering technique,” *IEEE Access*, vol. 8, pp. 83070–83080, 2020.
- [36] D. Tognetto, P. Cecchini, R. Giglio, and G. Turco, “Surface profiles of new generation intraocular lenses with improved intermediate vision,” *Journal of Cataract & Refractive Surgery*, vol. 46, no. 6, p. 1, 2020.
- [37] S. Q. Gu, Y. M. Ma, Z. K. Pan, and W. Z. Bo, “A variational level set model for implicit surface multiphase image segmentation,” *Computer Simulation*, vol. 37, no. 3, pp. 211–217, 2020.

Research Article

Discrete Dynamic Modeling of Learner Behavior Analysis in Physical Education Teaching

Jia Shi , Jun Sun , and Zhonghua Zheng 

Department of Sports, Central China Normal University, Wuhan, Hubei 430070, China

Correspondence should be addressed to Jun Sun; junsun@ccnu.edu.cn

Received 11 February 2022; Revised 17 March 2022; Accepted 21 March 2022; Published 12 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Jia 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.

With the advent of the big data era, the combination of information technology and education has become an important way for the development of the industry. The large-scale realization of teaching tasks under the background of information data requires the prediction and analysis of learners' characteristics, behavior, and development trend. Based on the above situation, this paper uses discrete dynamic modeling technology in big data environment to study the learners' behavior in physical education teaching. By quantifying the learning process data, the feature points of each learner are extracted to realize the personalized construction of dynamic learning data. Due to the rapid development of network technology, we mainly analyze the online education platform and explore the influencing factors of learners' behavior characteristics from many aspects. Finally, it carries out dynamic modeling and prediction for physical education learners from the aspect of achievement change, uses the grey model to build the achievement change system, and combines the dynamic modeling technology to reflect the development trend of achievement. The results show that the main factor affecting learners' behavior change in physical education is video learning. Most students are passive and lack initiative in the learning process. Discrete dynamic modeling technology can improve the accuracy of predicting student achievement changes and provide effective data for the research content.

1. Introduction

With the continuous development of the information age, the big data model has become a common data source in many fields (Lin et al., 2021) [1]. Information and data play an important role in human life. In the development of education, the traditional classroom teaching mode has gradually shifted to the modern teaching mode (Zhu et al., 2021) [2]. Using data information to solve the problems encountered in the educational process has become the norm (Cai and Lin, 2021) [3]. In the research of learner behavior analysis and learner feature extraction, the error impact caused by massive data is very obvious. We need to use a series of means to simplify data and extract hidden useful information from large-scale information (Yan et al., 2021) [4]. The ultimate goal is to help teachers accurately analyze students' learning and form targeted teaching methods. At the same time, it can also help learners improve their learning ability and change the traditional single

learning method (Shi and Hao, 2021) [5]. For the analysis of learners in the field of sports, it mainly explores from the aspects of intelligence, personalization, and online education. Due to the particularity of sports, the teaching environment needs suitable places as support. Therefore, many problems have been encountered in the transformation of education model (Liao et al., 2021) [6].

For the transformation of physical education teaching mode, we have added many products under big data technology to our teaching methods, using virtual reality technology as an implementation way of online integrated teaching (Zhang and Ge, 2021) [7]. Using virtual reality products in network teaching allows students to experience the immersive teaching environment. This way can improve students' interactive ability and information exchange ability in sports activities (Wang 2020) [8], realizing three-dimensional physical education through touch, hearing, vision, and other senses. Finally, the online learning students are analyzed for learner behavior, mainly using discrete

dynamic modeling technology in the context of big data (Zeng et al., 2020) [9]. This technology is used to solve the errors and storage difficulties caused by dynamic growth in the construction of data integration. Learner behavior analysis is the key content of educational research. The traditional research methods mainly use neural network algorithm and deep learning algorithm. Most of this research technology is in the static data environment and is studied from a certain perspective (He and Li, 2020) [10]. In fact, the changes of students' characteristics and behavior are in a dynamic mode, and the data changes are uncertain and dynamic as a whole. Therefore, in the analysis of learners' behavior, we mainly use discrete dynamic modeling technology to explore. We control students' learning styles and rules and study the factors affecting learners' behavior changes from many aspects. Finally, the overall performance of learners is dynamically predicted to obtain the optimal education model.

This paper analyzes the changing trend of learners' behavior in physical education teaching and its influencing factors on performance. This paper mainly discusses the application of data mining in learner behavior analysis and dynamic data processing. Innovative contributions include extracting students' key feature points by quantifying learning process data and providing learners with personalized learning strategies. Compared with the traditional learner behavior analysis model, discrete dynamic modeling can deal with the dynamic information in different environments and improve the overall accuracy of the model. Using the dynamic modeling method, this paper constructs the development and change model of sports performance and tests the results of the model. The error coefficient and accuracy of students' performance change are analyzed in the model. The results show that the discrete dynamic modeling of learner behavior analysis can help teachers effectively obtain students' learning state and learning effect.

This paper is mainly divided into three parts. The first part briefly describes the research status of learner behavior analysis and feature analysis, as well as the application background of discrete dynamic modeling technology. In the second part, firstly, the data mining algorithm is used to obtain student behavior data, and the differences between traditional modeling technology and discrete dynamic modeling technology are compared. Discrete dynamic modeling technology is used to analyze the behavior changes of physical education learners and build an intelligent learning guidance system. Finally, the grey model of learners' performance data in physical education teaching is constructed, and the discrete dynamic modeling is used to predict the change trend of performance. The third part first analyzes the results of the research on the behavior analysis and modeling of sports learners and analyzes the factors affecting the change of learners' behavior. Finally, the results of the prediction model of physical education learners' performance are analyzed.

2. Related Work

The research on learner behavior analysis is mainly applied to classroom teaching. Firstly, collect the behavior data generated by students in the learning process, compare the

teaching methods, and analyze the interaction between the two (Bi and Zhao, 2015) [11]. The obtained data will be processed and finally fed back to the classroom. Teachers improve teaching methods and optimize students' learning atmosphere by analyzing the results. In order to effectively predict the learning effect, many educational researchers extract the key feature point data according to the learning behavior pattern and construct the prediction model combined with the Bayesian algorithm in discrete dynamic modeling. The experiment mainly verifies the notion that learner behavior analysis can have good prediction effect. In Bayesian algorithm, the weight coefficient and correlation analysis are used to construct the prediction model, identify learners' status, and analyze the main factors affecting performance changes. The main purpose of studying the changes of learners' behavior in teaching in colleges and universities is to analyze students' learning status and carry out personalized teaching setting and intelligent guidance (Mu, 2021) [12].

Most colleges and universities in the United States collect the information data generated in the learning process, mainly to obtain teachers' teaching data and students' feedback data (Wang et al., 2021) [13]. The above information is used to help teachers optimize their teaching methods. By analyzing students' performance score, learning progress, learning effect, and learning style characteristics, this paper predicts whether the learners' whole learning environment is in a normal state. This kind of learner behavior analysis can help teachers quickly solve the problem of students' slack and obtain negative students' personal data. Through information processing, the corresponding intelligent learning methods are matched, combined with targeted optimization of teaching, to improve students' learning experience.

British universities began to monitor and study students' learning process at the end of the nineteenth century. At that time, the technical conditions were poor, and the analysis did not achieve good results (Deepak et al., 2019) [14]. With the rapid development of the information age, British education researchers use data mining and dynamic modeling technology to obtain and process students' learning data. It also makes an effective prediction for the trend of academic achievement. Finally, the analyzed data are fed back to the teaching process, which changes the traditional education model.

The research on online learning behavior patterns in Germany mainly predicts students' performance through identification, classification, diagnosis, and analysis (AbdEllatif et al., 2018) [15]. In the learner behavior model, the education system is constructed by extracting the student behavior sequence, which provides help for the design of the teaching system, using mobile data such as learning duration to analyze students' learning status, compare key behavior sequences, and accurately analyze the factors affecting students' performance. Based on the current situation of learner behavior analysis in the above countries, this paper mainly analyzes the physical education model. This paper studies the factors affecting the change of students' performance in physical education teaching, uses discrete

dynamic modeling to construct learners' behavior analysis system, and explores the state change of students in physical education teaching.

3. Methodology

3.1. Research on Discrete Dynamic Modeling of Learner Behavior Analysis in Physical Education Platform. With the development of computer science, pulse technology, microprocessor, and digital element components, a large number of systems in biology, ecology, aviation, aerospace, economy, and engineering control need to be described by discrete-time systems. When the continuous time system is analyzed and simulated by computer, it also needs to be discretized and then processed. With the expansion of the research field of system theory and the wide popularization and application of computer technology, discrete control system theory has developed rapidly and become an important part of control theory. Discrete system theory plays an important role in the fields of automatic control engineering, communication, radar technology, biology, power system, and nuclear physics. There are essential differences in the description of discrete systems and continuous systems, but there are similarities in analysis and research. Many concepts and research methods in continuous time systems can also be extended and applied to discrete systems.

With the exploration of students' learning analysis technology, more and more educational researchers begin to pay attention to the changes of learners' behavior. Many researchers have proposed a new way to explore learning behavior, that is, to transform and mine the behavior between learning associations. The sequence of behavior characteristics constructs a stable behavior analysis system. The learner behavior analysis system can reflect the time sequence of students' learning behavior, reflect the characteristics between individual and whole learners from a wide-angle perspective, and help teachers to effectively diagnose and improve the differences between different students. In physical education, we need to pay attention to the changes of learners' dynamic operation sequence and mine and extract students' feature points. Finally, the feature point information is identified for online learning pattern simulation, and the system model is constructed according to the original dynamic behavior sequence. The behavior analysis of physical education learning can provide data support for students' learning environment and final performance prediction. This paper constructs a system framework for the online learning behavior model, mainly including the behavior sequence of goals in sports, students' key action events, and behavior changes. Descriptive means are used to identify students' behavior characteristics and complete the preliminary construction of the system model. The framework of learner behavior analysis model is shown in Figure 1.

As can be seen from Figure 1, the behavior mode is mainly composed of description analysis, diagnosis result analysis, and prediction analysis. Through the exploration of all aspects of students, a circular model is formed, extracting the characteristics of target activities from the process of

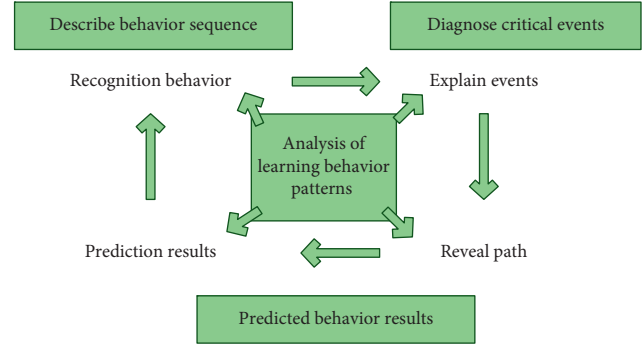


FIGURE 1: Learner behavior analysis model framework.

target activities is the basis for analyzing target activities. Firstly, the logical regression algorithm is used to analyze and process the behavior data generated by learners, and the function variables and data loss variables are constructed:

$$P(y = 1|x) = \sigma(wx + b) = \frac{e^{wx+b}}{e^{wx+b} + 1}, \quad (1)$$

$$P(y = 0|x) = 1 - \sigma(wx + b) = \frac{1}{e^{wx+b} + 1}.$$

The above formula can obtain the function value after data processing and the data value lost in the analysis process. When exploring the correlation between learning behavior and result analysis, we use the correlation coefficient of linear regression. The calculation method is as follows:

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

Among them, the value range of correlation coefficient r is between positive and negative 1. When the function value is positive, the two variables have correlation. In physical education, the way to study the change of behavior sequence should be analyzed from the perspective of students' movement change. Through the state of students' learning process and the behavior data generated by the online platform, the frequency of students' behavior and activity in different age groups were calculated. Secondly, all learners are classified and analyzed according to the change of feature sequence. Classification and characterization can be used to show the particularity between different groups. Dynamic changes often occur in the processing of students' personal data, for example, replacing and deleting personal data. This situation will also lead to the repetition of historical traces and affect the accuracy of prediction results. Therefore, this paper proposes discrete dynamic modeling technology as a method to build system model. Learner behavior analysis modeling is the process of realizing personalized teaching. The performance of each student in different environments also determines the trend of the overall data. The key feature points of learners are extracted by quantifying student data, and finally the construction of dynamic modeling is realized. Since this process is a random discrete dynamic state, in

order to summarize most of the key characteristics and time series changes, this paper adopts a three-stage dynamic modeling path, and the specific process is shown in Figure 2.

As can be seen from Figure 2, the research expresses the behavior analysis and modeling path of physical education learners, which is completed in three stages: quantification of learning data, feature analysis and extraction, and finally recommendation of personality learning methods. Each student will be affected by many factors in different environments, such as learning ability, cognitive level, learning efficiency, and adaptability. In order to analyze the time dimension of learners' behavior change, we use Bayesian network for evolutionary analysis in various places of physical education teaching. The learning process of middle school students in physical education teaching environment is dynamically modeled and evolved, and the personality feature points are extracted. It has made an effective contribution to improving the learning efficiency of middle school students.

Finally, the student feature points after discrete dynamic modeling and analysis are uniformly classified, the repeated links in the data are simplified, and an intelligent guidance system for physical education teaching is formed. Through the analysis of physical education learners' behavior, intelligent guidance can provide targeted learning strategies for each student, promoting students to obtain suitable learning resources in specific ways. With the support of the online platform, learning resources have also been expanded to images, videos, virtual reality scenes, and so on. Because there are spatial attributes such as sparsity and structure between student behaviors, it is necessary to simplify the combination of feature points of data variables. Therefore, discrete dynamic modeling can realize the data processing function under multiangle correlation. For students with different learning environments and learning methods, we comprehensively explore their learning time, learning speed, and learning methods. Combined with discrete dynamic modeling analysis, a behavior inquiry model of physical education learners is formed. It can realize personalized learning strategy recommendation for students from various aspects and automatically make dynamic adjustment according to learners' feedback information.

3.2. Research on Student Achievement Prediction Modeling under the Analysis of Learners' Behavior Data in Physical Education Teaching. In physical education teaching activities, the time series prediction method is basically used to predict the change of individual sports performance. According to the time-varying growth coefficient of activity scores as data, mathematical modeling is selected for trend analysis. This situation can easily lead to the expansion of the factors affecting students' sports performance, and the result data is easy to be unknown and fuzzy. Therefore, in the change of learners' behavior in physical education teaching, the prediction of achievement first selects the mathematical model for analysis. This paper focuses on grey model

prediction, combined with discrete dynamic modeling technology, to explore the changes of physical education learners' behavior and achievement.

The content of grey model theory is that whether the data is complex or not, it is interrelated, orderly, and in the overall scope. The characteristics of the data cannot use the original sequence model, because the original prediction method cannot accurately calculate the processed data information for the irregular sports results. Firstly, the mined data is used to process the score series in order to provide basic data for subsequent discrete modeling. The set variables are defined as follows:

$$Y_0; Y_0(1), \dots, Y_0(i), \dots \quad (3)$$

The above formula is used to represent the time-varying performance series. The new sequence formed by accumulating the data is

$$Y_1(1) = \sum_{i=1}^i Y_0(i) \quad (4)$$

where $Y_0(i)$ is the definition of continuous change in the data sequence formed by accumulation. According to the calculation results, we find that the data show discrete changes, and the generated weighted values rise monotonically. The overall randomness of the data has decreased, which is convenient for us to optimize the accuracy of student performance prediction. General models are measured in stages. For the change of sports performance, we set the corresponding generated value about single data. Further considering the discrete change of numerical value, the performance trend is basically in a stable state. This model is a differential calculation equation:

$$\frac{dY_1(t)}{dt} - KY_1(t) = \beta. \quad (5)$$

The above formula is equivalent to

$$\frac{dY_1 d/dt}{Y_1 d} = K + \frac{\beta}{Y_1 d}. \quad (6)$$

Among them, K and β are the determined values in the achievement sequence vector in sports, and K is the growth efficiency in the cumulative data. We can customize the initial conditions and get the solution of the equation as follows:

$$\hat{Y}_1(t) = \left(Y_0(1) + \frac{\beta}{K} \right) e^{K(t-1)} - \frac{\beta}{K}. \quad (7)$$

The final equation is obtained by calculating the score sequence of a student. The minimum solution of the equation is

$$\begin{pmatrix} \hat{K} \\ \hat{\beta} \end{pmatrix} = (B^T B)^{-1} B^T X, \quad (8)$$

in which

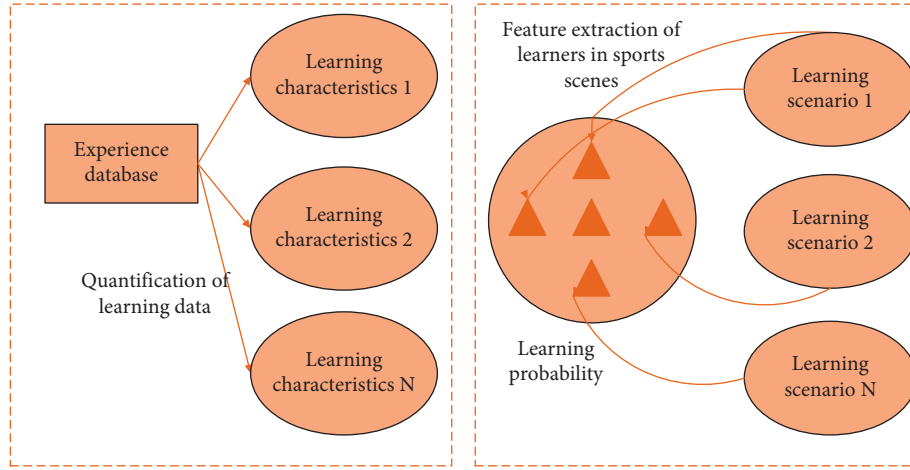


FIGURE 2: Dynamic modeling path.

$$B = \begin{bmatrix} \frac{1}{2} [Y_1(1) + Y_t(2)] \\ \frac{1}{2} [Y_1(2) + Y_t(3)] \\ \dots \\ \frac{1}{2} [Y_1(N-1) + Y_t(N)] \end{bmatrix}. \quad (9)$$

The variable optimal solution of the equation is

$$\lambda = (Y_0(2), Y_0(3), \dots, Y_0(N))^T, \quad (10)$$

$$Y_1(i) = Y_1(i-1) + Y_0(i)^i = 2, 3, \dots, N.$$

Discretize the above formula, bring in the defined variables, and obtain the performance development trend model of physical education learners:

$$\hat{Y}_1(i+1) = \left(Y_0(1) + \frac{\hat{B}}{\hat{K}} \right) e^{Ki} - \frac{\hat{B}}{\hat{K}} \quad i = 1, 2, \dots \quad (11)$$

According to the formula, it is also necessary to restore the performance prediction process, that is, to obtain the value of the prediction time series model:

$$\hat{Y}_1(i) = \left(Y_0(1) + \frac{\hat{B}}{\hat{K}} \right) (e^{Ki} - e^{K(i-1)}) \quad i = 1, 2, \dots \quad (12)$$

Finally, test the results of the grey model of physical education learners' performance, and set the relevant error coefficient:

$$\delta(i) = Y_0(i) - \hat{Y}_0(i) \quad i = 1, \dots, N. \quad (13)$$

This original grey model is commonly used by most researchers to predict learners' achievement in physical education teaching. On this basis, we combine discrete dynamic modeling technology for optimization. Through the calculation process of the above formula, we find that there are much discrete data in the calculation, which is very

helpful to our dynamic modeling. In the discrete dynamic modeling model, we store individual data as subsets in the data system, and each subset will be used many times in processing. Finally, the data effect is analyzed in different cases to obtain the optimal output results. That is to analyze and process the learner's behavior data for many times and select the value with the least repeatability and the least storage for application. In this paper, the accuracy difference between the ordinary grey model and the model optimized by discrete dynamic modeling in the prediction of students' achievement change in physical education teaching is shown in Figure 3.

As can be seen from Figure 3, in the model test, the accuracy of the ordinary grey model decreases significantly with the increase of learner behavior data. Compared with the average prediction accuracy, this paper uses discrete dynamic modeling technology, and the optimized model can maintain the basic accuracy of the model.

4. Result Analysis and Discussion

4.1. Discrete Dynamic Modeling of Learner Behavior Analysis on Physical Education Teaching Platform. In the analysis of learners' behavior, it mainly involves the following aspects: first, whether each student has prepared before physical education class. The second is the activity of students' online learning and whether they pay attention. Finally, it lies in the situation of watching PE teaching video and students' feedback on video teaching. On the online platform, we control the number of published learning tasks and explore the number of classroom tasks completed by each student. Learners can choose tasks independently according to their own learning needs. Because students' learning environment is different, each student's learning ability is also different, so the data we obtain is discrete and dynamic. The information data faced by traditional modeling technology is static, while discrete dynamic modeling can quickly process dynamic data for effective analysis. In this paper, students of the same grade are randomly selected as experimental variables, and student behavior data are taken as independent variables. The correctness of the efficiency test of traditional modeling

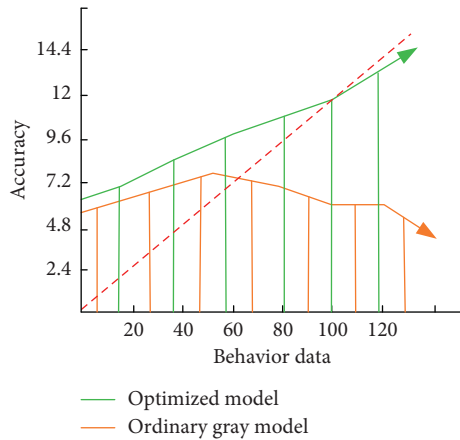


FIGURE 3: Comparison of accuracy between ordinary grey model and optimized model.

technology and discrete dynamic modeling technology for completing tasks is explored, as shown in Figure 4.

In the dynamic data of learners' behavior changes, discrete dynamic modeling technology can accurately grasp the efficiency changes of each student in completing tasks in class. As can be seen from Figure 4, compared with the efficiency of actually completing the task, the efficiency deviation of traditional modeling technology inspection is large. Therefore, in the specific experimental process, students' ability to master physical education teaching knowledge is different.

By setting the theoretical knowledge of topic detection, we can effectively judge students' learning. Before the formal class, according to the learner behavior analysis model, teachers can understand each student's advantages and knowledge points in advance. Therefore, we can judge whether students have a slack learning attitude in the whole classroom.

4.2. Analysis of Research Results of Student Achievement Prediction Modeling under the Analysis of Learner Behavior Data in Physical Education Teaching. The data used in the analysis of learners comes from the information produced by the learning process of students in physical education teaching. Data analysis mainly includes collection, acquisition, storage, representation, and other operations. We take the behavior data of college students in physical education classroom as the research content, explore the impact of learners' behavior data on students' performance, and finally predict the trend of students' performance. Through the analysis of students' physical activity records and final scores, teachers can understand the factors affecting students' learning effect in the classroom. Among them, data collection is the most basic part in the prediction process, and it is also the most important link of trend analysis. We use data standards for model construction and use discrete dynamic modeling to process dynamic data. This study obtained the effective class records of more than 9000 learners in the online platform. The data of volleyball, basketball, and tennis in physical education teaching are

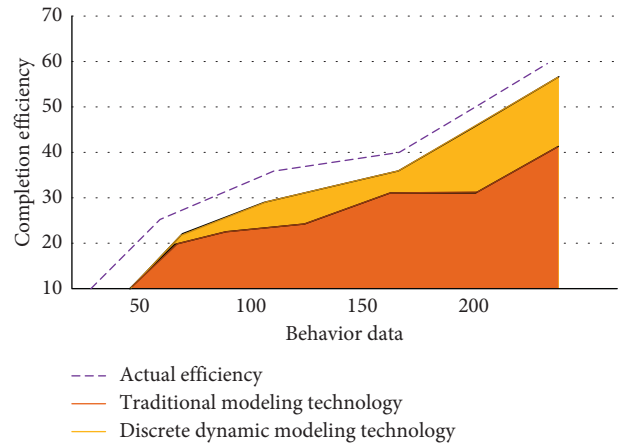


FIGURE 4: Comparison between traditional modeling technology and discrete dynamic modeling technology for the correctness of task completion efficiency test.

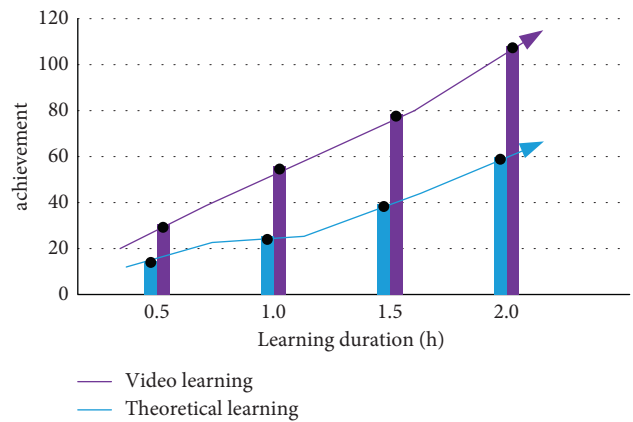


FIGURE 5: Change trend of students' performance under different modes.

classified. The specific calculation indicators need to consider the time and overall learning duration of students logging in to the learning platform. Before exploring the factors affecting students' performance, the chaotic and repeated information generated in the data processing process is filtered and screened by discrete dynamic modeling. We ensure that the data obtained is single, simplified, and highly related to physical education teaching. Finally, we need to clean up the data task and delete irrelevant features and recode them. This part of variable data has a great impact on performance prediction.

In the teaching method, we found that the change trend of students' performance through video learning is more obvious. Physical education teachers uploaded 98 video courses in the course, with a total length of 150 minutes. We take students' learning behavior as the monitoring object and select students from two classes to explore. Set the video learning control group. With the increase of learning time, the change of students' performance shows different trends, as shown in Figure 5.

As can be seen from Figure 5, with the increase of learning time, the upward trend of students' performance in

general theory teaching is not obvious. The completion rate of students learning through video is about more than 90%, and their grades are obviously rising. Therefore, we preliminarily infer that students' learning situation and enthusiasm are related to learning styles. Most students have high enthusiasm, but in physical education, video learning is more suitable.

5. Conclusion

In the big data environment, we use a variety of query methods to analyze the behavior characteristics of learners. Learners' learning ability, emotional attitude, and learning time will affect their performance. In the physical education teaching mode, because its own educational process is different from other disciplines, it needs a large number of practical courses and places as support. The differences of each student's physical quality and learning style will produce corresponding dynamic changes. This paper presents a prediction model based on discrete dynamic modeling technology. This paper analyzes the changing trend of learners' behavior in physical education teaching and its influencing factors on performance. This paper mainly discusses the application of data mining in learner behavior analysis and dynamic data processing. In the learning behavior pattern, feature serialization is used for data encoding. This coding method is a complex data processing process, so we need to carry out effective modeling and analysis after data mining. Secondly, by quantifying the learning process data, the key feature points of students are extracted to provide learners with personalized learning strategies. Compared with the traditional learner behavior analysis model, discrete dynamic modeling can deal with the dynamic information in different environments and improve the overall accuracy of the model. Finally, the dynamic modeling method is used to construct the development and change model of sports performance, and the results of the model are tested. However, the error coefficient and accuracy of the analysis model need to be further discussed. Future research needs to help teachers effectively obtain students' learning state and learning effect on the discrete dynamic modeling of learner behavior analysis.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] X. Lin, "Research on the application of K-means clustering algorithm in online learning behavior in the era of big data," *Electronic design engineering*, vol. 29, no. 18, pp. 181–184, 2021.
- [2] B. Zhu, tongna Shi, Z. Shi et al., "Construction of "pyramid" e-learning platform and analysis of learning behavior," *Experimental Technology and Management*, vol. 38, no. 8, pp. 208–212, 2021, + 216.
- [3] G. Tripathi, K. Singh, and D. K. Vishwakarma, "Convolutional neural networks for crowd behaviour analysis: a survey," *The Visual Computer*, vol. 35, no. 5, pp. 753–776, 2019.
- [4] A. Yan, Y. Hu, and X. Xing, "Research on online learning behavior characteristics of learners based on learning analysis technology," *Anhui Sports Science and technology*, vol. 42, no. 2, pp. 68–73, 2021.
- [5] N. Shi and X. Huo, "Research on the optimization of College Business English mixed curriculum system based on learning behavior analysis," *Journal of Heihe University*, vol. 12, no. 4, pp. 112–113, 2021.
- [6] Y. Liao, S. Yang, and Y. sun, "Analysis and Research on online learning behavior based on SPSS," *Modern information technology*, vol. 5, no. 7, pp. 21–24, 2021.
- [7] L. Zhang and F. Ge, "Research on the construction of learner behavior analysis platform based on big data," *China education informatization*, no. 5, pp. 50–53, 2021.
- [8] Bo Wang, "Research on online education learning behavior analysis technology based on big data," *Electronic world*, no. 24, pp. 100–101, 2020.
- [9] H. Zeng, Y. Zheng, and Y. Huang, "Analysis and Research on Teachers' online learning behavior based on the concept of big data," *Journal of Guangzhou Radio&TV University*, vol. 20, no. 6, pp. 29–33, 2020.
- [10] dianchen He and Li Li, "The influence of computer modeling in Physical Education [J]," *Journal of Shanxi Normal University (Philosophy and Social Sciences edition)*, vol. 38, no. 1, pp. 85–92, 2020.
- [11] mujin Bi and D. Zhao, "Teaching exploration of mathematical modeling thought in higher vocational colleges of physical education," *Sports*, no. 14, pp. 126–127, 2015.
- [12] X. Mu, "Practical path of mathematical modeling teaching reform based on project teaching method," *Journal of Xinxiang University*, vol. 38, no. 9, pp. 61–64, 2021.
- [13] G. Wang, Q. S. Jia, and M. C. Zhou, "Artificial neural networks for water quality soft-sensing in wastewater treatment: a review," *Artificial Intelligence Review*, vol. 55, pp. 1–23, 2021.
- [14] G. Deepak, A. Ahmed, and B. Skanda, "An intelligent inventive system for personalised webpage recommendation based on ontology semantics[J]," *International Journal of Intelligent Systems Technologies and Applications*, vol. 18, no. 1-2, pp. 115–132, 2019.
- [15] M. AbdEllatif, M. S. Farhan, and N. S. Shehata, "Overcoming business process reengineering obstacles using ontology-based knowledge map methodology," *Future Computing and Informatics Journal*, vol. 3, no. 1, pp. 7–28, 2018.

Research Article

SCU-Net: Semantic Segmentation Network for Learning Channel Information on Remote Sensing Images

Wei Wang ¹, Yuxi Kang ¹, Guanqun Liu ², and Xin Wang ¹

¹School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410114, China

²Hunan Open University, Changsha 410004, China

Correspondence should be addressed to Guanqun Liu; patrice@126.com and Xin Wang; wangxin@csust.edu.cn

Received 21 January 2022; Accepted 9 March 2022; Published 10 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Wei 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.

Extracting detailed information from remote sensing images is an important direction in semantic segmentation. Not only the amounts of parameters and calculations of the network model in the learning process but also the prediction effect after learning must be considered. This paper designs a new module, the upsampling convolution-deconvolution module (CDeConv). On the basis of CDeConv, a convolutional neural network (CNN) with a channel attention mechanism for semantic segmentation is proposed as a channel upsampling network (SCU-Net). SCU-Net has been verified by experiments. The mean intersection-over-union (MIOU) of the SCU-Net-102-A model reaches 55.84%, the pixel accuracy is 91.53%, and the frequency weighted intersection-over-union (FWIU) is 85.83%. Compared with some of the state-of-the-art methods, SCU-Net can learn more detailed information in the channel and has better generalization capabilities.

1. Introduction

Remote sensing image processing technology has developed rapidly. The image semantic segmentation approach based on CNN has become a hot research direction in satellite remote sensing image semantic segmentation [1], and it is widely used in land detection, vegetation classification, environmental monitoring, urban planning, national defense security, and other fields. At the same time, the neural network can automatically obtain the information contained in the pictures and accurately predict the category of unknown data [2]. Therefore, academia began to use neural networks to obtain the deep features in high-resolution remote sensing images and use them for satellite remote sensing image classification [3].

Recently, the deep learning method has performed well in the computer vision recognition task, and CNN has achieved very good results in image classification [4–9], semantic segmentation [10], and target detection [11, 12]. The realization of LeNet [13] established the structure of CNN, and classic CNNs, such as VGG [14], GoogleNet [15], and ResNet [16] have made great achievements. FC-

DenseNet (fully convolutional DenseNet) was proposed by Simon et al. [17] to enhance feature extraction and feature reuse. Based on a fully convolutional network, the standard convolutional layer is replaced by dense block [18], and the resulting network can achieve new pixel-level segmentation tasks rather than traditional image classification tasks. To make the output size of the network equal to the input size, the upsampling structure of FCN is added. The network achieves a good segmentation effect in semantic segmentation. The DFCN (dense fully convolutional network) proposed by Wang et al. [19], which also combined DenseNet and fully convolutional network, introduced the dense blocks in the dense network into the full convolutional network. Finally, a good segmentation effect has been achieved on the satellite remote sensing image dataset.

There are many important methods of image processing, among which image semantic segmentation is one. Semantic segmentation is to divide the pixels in the image and predict the category of each pixel in the image using the neural network [20], which is like image super-resolution reconstruction [21]. Recently, geographic

information system (GIS), self-driving car, medical image analysis, and robot have become the main application fields of semantic segmentation. This paper mainly studies the application of semantic segmentation methods in the field of geographic information systems. There have been many pieces of research on semantic segmentation using deep learning methods in recent years. The development and application of CNN [5] are introduced by Long et al. In 2015, the FCN for image semantic segmentation was proposed by Long [10]. It adjusts the structure of the ordinary convolution network and realizes the intensive prediction of the image by abandoning the fully connected layer. At the same time, a skip connection is added to the network. This method can not only extract the rough semantic information from the deep layer but also combine the appearance information from the shallow layer, which makes the segmentation more refined. The FCN convolution neural network is used to realize the leap from image pixels to pixel categories. For the medical image segmentation problem, inspired by the FCN network, Fischer et al. proposed U-Net [22]. Through the design of a U-shaped network structure in U-net, not only context information but also location information can be obtained, which has achieved good results in biomedical segmentation applications. U-Net won the first prize for outstanding performance in the 2015 ISBI cell tracking competition. SegNet [23] was proposed by Badrinarayanan et al. It is based on the encoder-decoder structure, which reduces the number of network parameters by discarding the location information in the pooling layer during the upsampling process and achieves good results. The researchers conducted experiments on two datasets, one is the SUN RGB-D dataset used for indoor scene segmentation. The other is the use of the CamVid dataset for road scene segmentation. The DeepLab series network was proposed by Google. DeepLabV1 [24] is an improved semantic segmentation network based on VGG16. Researchers found in experiments that the accuracy of using deep convolutional neural networks in the field of semantic segmentation is insufficient. The translation invariance of the advanced features of deep convolutional neural networks is the fundamental reason. DeepLabv1 combines a fully connected conditional random field (CRF) with the response of a deep convolutional neural network to try to solve this problem. Then, the Hole (dilated convolution) algorithm is creatively integrated into the deep convolution neural network model, which performs well on the PASCAL VOC 2012 dataset. An image segmentation algorithm based on spatial pyramid pool (ASPP) structure was proposed by DeepLabV2 [25]. ASPP can capture the contextual content of objects and images at multiple scales and detect incoming convolution features using convolution kernels with multiple sampling rates and effective field of view. DeepLabv2 further highlights the role of hole convolution in dense prediction tasks. In the deep convolution network, the dilated convolution can effectively control the resolution (receptive field) of calculating the feature responses to effectively expand the receptive field of convolution kernels

without additional parameters or calculation. Integrate more context information. DeepLabV3 [26] uses the multigrid strategy and adds batch normalization, global pooling layer, point convolution layer, and bilinear interpolation upsampling to the ASPP module, achieving good results. The DeepLabV3+ [27] network uses the Xception [28] network as the backbone and uses the encoder-decoder structure. In addition, both [29, 30] use support vector machine (SVM) to segment remote sensing images.

Inspired by these studies, this paper proposes a new method for semantic segmentation of satellite remote sensing images. In the task of semantic segmentation of satellite remote sensing images, the improved upsampling method proposed in this paper not only reduces the number of model parameters but also improves the recognition ability of the model. Based on the new upsampling module, a novel image semantic segmentation model and introduced attention mechanism are proposed. In addition, a new remote sensing image dataset is provided, which shows the remote sensing information of Chenzhou, China. These satellite remote sensing images come from China's GF-2 satellite with a spatial resolution of 0.8 meters, showing various spatial information of cultivated land, forest land, waters, buildings, etc. The original images in the dataset are labeled based on local geomorphological features. Finally, the method in this article was tested on this dataset. Compared with other semantic segmentation methods based on deep learning, SCU-Net has better segmentation performance and stronger generalization ability.

2. SCU-Net

This section will, firstly, introduce the structure of FCN. Secondly, the structure and principle of SCU-Net proposed in this paper are introduced. Finally, the network complexity of SCU-Net is analyzed through experiments.

2.1. Fully Convolutional Network. The main difference between the FCN network and the traditional deep learning convolutional network is that it does not use a full connection layer. The FCN network adds an upsampling method to the network structure. The output image has the same size as the original input image, which realizes pixel classification. FCN is mainly composed of the feature extraction module of the ordinary convolutional network, upsampling, and skip connection. Figure 1 shows the structure of FCN, where "Input" represents the input end of the image, and "Conv1," "Conv2," "Conv3," "Conv4," and "Conv5" represent ordinary convolutional layers. "Max-Pooling" means the pooling layer. "Upsampling-1" means 2 times the upsampling module, "Upsampling-2" means 4 times the upsampling module, "Upsampling-3" means 8 times the upsampling module, and three upsampling modules can process the feature map to 2 times, 4 times, and 8 times the size of the original image.

The first operation is feature extraction. If the input image size is $H \times W$, the original image will be reduced to

$(H \times W)/2$ after the “Conv1” and “MaxPooling” operations. Then, the feature map is reduced to $(H \times W)/4$ through the operations of “conv2” and “MaxPooling.” By using the “Conv3” and “MaxPooling” operations, the output feature map “featureMap-3” with reduced size $(H \times W)/8$ can be obtained. Then, “Conv4” and “MaxPooling” operations are performed to reduce the size of the feature map to $(H \times W)/16$, and the output feature map is named “featureMap-4.” In the next step, the feature map goes through “Conv5” and the last “MaxPooling” operation, and the feature map, namely “featureMap-5,” will be reduced to $(H \times W)/32$.

After the feature extraction is completed, the upsampling operation is performed. First, the “Upsampling-1” operation is performed on “featureMap-4” to obtain a feature map “featureMap-6” with a size of $(H \times W)/8$. Next, “featureMap-5” performs the “Upsampling-2” operation to obtain the feature map “featureMap-7” whose size is $(H \times W)/8$. Then, “featureMap-6” and “featureMap-7” are stitched together. Finally, the “Upsampling-3” operation is performed on the output result to output the feature map. The size of this feature map is the same as the original image, and the semantic segmentation is completed.

2.2. Network Structure. The channel feature weight extraction module (CFWE) [7] is introduced to improve the feature extraction ability of the semantic segmentation network model. Figure 2 is the detailed structure of CFWE. In this structural diagram, “Conv k ” denotes a convolutional layer with a filter size of k , where k is equal to 3 or 1. “GAP” denotes the global average pooling layer, and “FC” denotes the full connected layer.

The CFWE module is a multilayer structure. The first is the short connections layer of the CFWE. It contains 3 convolutional layers, of which, there are two conv1 and one conv3. The two “conv1” in the module are used to increase and decrease the number of channels, respectively, with the purpose of reducing the amounts of parameters. The shortcut connections layer can alleviate the network degradation problem to a certain extent. Next comes the serial pool layer and full connection layer structure, which consists of two “FC” structures and a “GAP” structure. Firstly, the feature map of each channel is processed as a global feature using the “GAP” structure. Then, the feature dimension is reduced through the first “FC” layer, and the feature dimension is restored by the second “FC” layer. The weight coefficient model of each channel can be learned after passing through two “FC” layers. Finally, the final output feature map is to multiply the original feature and the extracted feature map channel weight coefficients by short connections. When extracting features, the weight coefficient is useful for the model to extract more important channel features and then enhance the feature extraction ability of the network.

In the traditional semantic segmentation neural network, the upsampling step is divided into two steps. Firstly, the feature map size is up-sampled to the original image size, and then point convolution is performed to complete the semantic segmentation. The parameters and calculations of

the network model using this method are relatively large. The “CDeConv” module proposed, firstly, performs point convolution to make the number of feature map channels consistent with the number of segmentation task categories, and then, it performs deconvolution to make the output feature map size consistent with the original image size. This module not only reduces the number of model calculations and parameters but also improves model performance. Its structure is shown in Figure 3.

According to the CFWE module and the “CDeConv” module, a novel semantic segmentation network model SCU-Net is proposed, whose structure is shown in Figure 4. Among them, “Input” represents the input image, and “Output” represents the output result. “Conv” not only includes “Convolution,” but also “Batch Normalization” and “Activation Function.” The function of this composite structure is to extract image features. “Conv7” is a “Conv” structure with k equal to 7 and a stride equal to 2. “Conv1” is a “Conv” structure with k equal to 1 and a stride equal to 1. “Conv1-3-1” is a structure consisting of “Conv1,” “Conv3,” and “Conv1” in series. One of the two “Conv1” structures is used to increase the dimensionality and the other can decrease the dimensionality. This operation can reduce the complexity of the model. “Conv3” is used for feature extraction. “Maxpooling” means the maximum pooling layer with a step size of 2 and a window size of 3. Its main function is to upsample, reduce dimensionality, and reduce the amount of calculation. “Upsampling” represents an 8 times upsampling module, and its function is to upsample the feature map to the same size as the original image to achieve semantic segmentation.

If the input image is X and the size is $H \times W \times 3$, it, firstly, passes through the “Conv7” module with 64 convolution kernels. The step size is 2, and the convolution kernel size is 7. Set the output feature map of the “Conv7” module to X_1 . The size becomes $H/2 \times W/2 \times 64$, which is shown as follows:

$$X_1 = \text{Conv7}(X). \quad (1)$$

After a maxpool layer, the number of channels will not change, however, the size of the output feature map X_2 will become 0.25 times that of X_1 , as shown below.

$$X_2 = \text{MaxPooling}(X_1). \quad (2)$$

Then, through n (n is a settable parameter, related to the number of network layers) serially connected “Conv1_3_1,” the output size of the feature map is $H/4 \times W/4 \times 256$. Next, the output X_3 can be obtained by using CFWE module. Its feature map size is $H/4 \times W/4 \times 256$, which can be expressed as follows:

$$X_3 = \text{CWFE}(\text{Conv1_3_1}(\text{Conv1_3_1}(\dots(X_2)))). \quad (3)$$

After that, the feature map X_3 is input to the n series of “Conv1_3_1” modules, and then through a layer of CFWE module, the output feature map X_4 has a size of $H/8 \times W/8 \times 512$, which can be expressed as follows:

$$X_4 = \text{CWFE}(\text{Conv1_3_1}(\text{Conv1_3_1}(\dots(X_3)))). \quad (4)$$

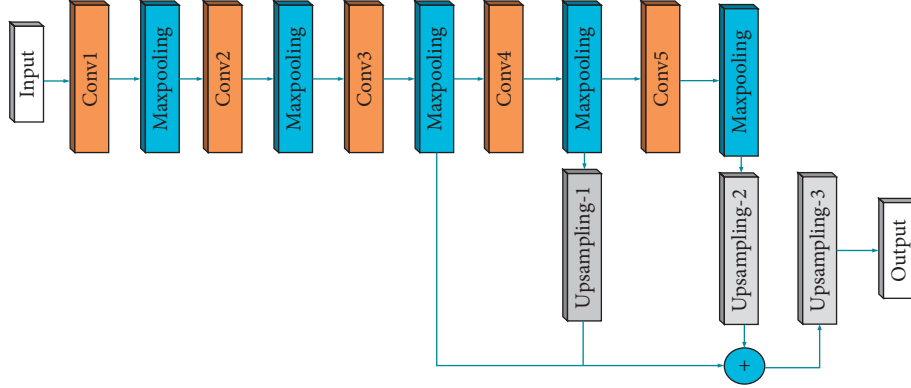


FIGURE 1: The structure of FCN.

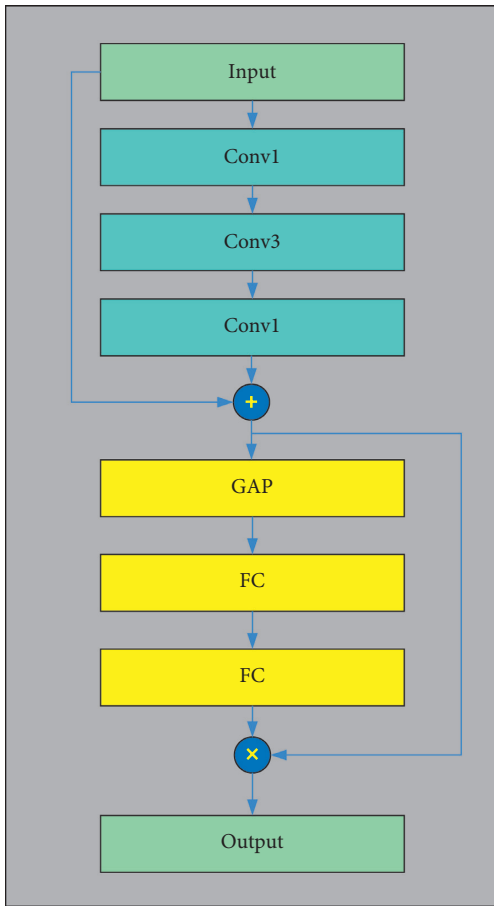


FIGURE 2: The structure of CFWE.

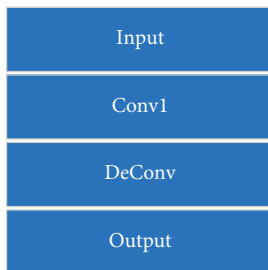


FIGURE 3: The structure of CDeConv.

The feature map X_4 is input to the n series of “Conv1_3_1” modules, and then through a layer of CFWE module, the output feature map X_5 size is $H/16 \times W/16 \times 1024$, which can be expressed as follows:

$$X_5 = \text{CWFE}(\text{Conv1_3_1}(\text{Conv1_3_1}(\dots(X_4)))). \quad (5)$$

Then, the feature map X_5 is input to n serially connected “Conv1_3_1” modules, and the output feature map X_6 has a size of $H/32 \times W/32 \times 2048$, which can be expressed as follows:

$$X_6 = \text{CWFE}(\text{Conv1_3_1}(\text{Conv1_3_1}(\dots(X_5)))). \quad (6)$$

The feature map X_6 uses a CDeConv module to perform point convolution, firstly, to achieve the purpose of reducing the number of channels, reducing the number of channels to 7. Then, upsample 4 times, and get the output X_7 . Its feature map size is $H/8 \times W/8 \times 7$, which can be expressed as follows:

$$X_7 = \text{CDeConv}(X_6). \quad (7)$$

Feature map X_5 enters a CDeConv module through a skip connection. Firstly, point convolution is performed to reduce the number of channels to 7. Then, upsample 2 times. Get the output X_7' . Its feature map size is $H/8 \times W/8 \times 7$, which can be represented as follows:

$$X_7' = \text{CDeConv}(X_5). \quad (8)$$

The feature map X_4 reduces the number of channels to 7 through a point convolution, and the size of the output feature map X_7'' is $H/8 \times W/8 \times 7$, which can be represented as follows:

$$X_7'' = \text{CDeConv}(X_4). \quad (9)$$

Then, add X_7 , X_7' , and X_7'' bit-by-bit to get X_8 . Its feature map size is $H/8 \times W/8 \times 7$, which can be expressed as follows:

$$X_8 = X_7 \oplus X_7' \oplus X_7''. \quad (10)$$

The final feature map X_8 passes an 8 times upsampling module to output the final output feature map, whose size is $H \times W \times 7$, and can be expressed as

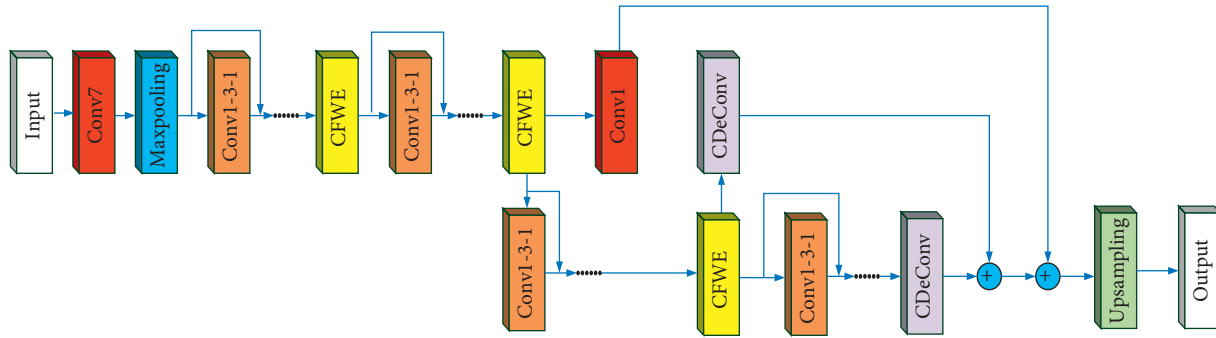


FIGURE 4: The structure of SCU-Net.

Three different upsampling methods using different numbers of “CDeConv” modules and skip connections are designed. Figure 4 shows SCU-Net-A. Based on SCU-Net-A, the network obtained by deleting the first “Conv1” module and a corresponding skip connection on the network backbone is SCU-Net-B. Based on SCU-Net-B, the network obtained by removing a “CDeConv” module connected to the CFWE module is SCU-Net-C.

Based on SCU-Net, combined with data expansion technology for image preprocessing, an automatic remote sensing image segmentation method is proposed. Figure 5 shows the specific process of the network model. Firstly, preprocess the remote sensing image data, and then it is input into SCU-Net for image semantic segmentation, and finally, the network outputs the segmented image.

2.3. Network Complexity. Model complexity refers to the amount of calculation and parameters of the network. The parameters and calculations of different upsampling methods and depth models are compared through the experiments. Experiments were performed with one “CDeConv,” two “CDeConv,” and three “CDeConv” modules in the network model using three different upsampling methods. By combining three different upsampling methods and three different network depths, we can get 9 different network models. Figure 6 is a comparison diagram of parameter quantities. Figure 7 is a comparison chart of the number of calculations.

As shown in Figure 6, when the depth of the SCU-Net is the same, the parameters of the C structure are about 10,000 less than the network parameters of the A and B structures. It shows that different upsampling structures have little influence on the parameters of the network model. The parameter quantity of SCU-Net-153-A is 2.46 times that of SCU-Net-51-A, and the parameter quantity of SCU-Net-102-A is 1.80 times that of SCU-Net-51-A. Therefore, it can be inferred that the number of model parameters is most affected by the number of network layers. Therefore, when the equipment does not support a large parameter scale, it is better to use a shallow depth network model.

The floating point of operations refers to the number of multiplication and addition operations in the model. As shown in Figure 7, the depth of the network has a great

influence on the amount of calculation. As far as the amount of calculation is concerned, SCU-Net-153 is 1.44 times that of SCU-Net-102, and SCU-Net-102 is 1.78 times that of SCU-Net-51. The calculation amount of SCU-Net-153 and SCU-Net-102 is very huge. Therefore, when the model accuracy gap is small, the SCU-Net-51 model has the lowest cost.

3. Experimental Results

3.1. Dataset. This experimental dataset comes from the fusion image of GF-2 in the Chenzhou area in 2016. As shown in Figure 8, there are satellite remote sensing images in the dataset. The landform of the Chenzhou area is complex, and the forest is rich in plant species. The number of wavebands is 3, and the spatial resolution of this remote sensing image is 0.8 meters. The original Gaofen-2 satellite images are preprocessed, cropped, annotated in the range of 2000×2000 pixels, and annotated various image types into ground truth images with different colors. According to the landform characteristics of Chenzhou, the species are divided into seven categories: cultivated land, road, forest land, buildings, water area, ridge and ditch, and others. The dataset contains 12,000 images, and the length and width of each image are 256×256 . Randomly select 10,000 pictures as the training set and the remaining 2000 pictures as the test set. Figure 9 shows the partial data cut into a uniform size from a large remote sensing image.

3.2. Preprocessing and Experiment Setup. In the deep learning network, it is easy to cause memory overflow when large remote sensing images to be classified are directly input into the network model. It is difficult for the model to identify the original data without preprocessing. Furthermore, training a deep learning network model needs lots of training data, however, labeling geographic information in satellite remote sensing images is a very complicated and time-consuming task, and labor costs are too high, making it difficult to produce large datasets. Additionally, for avoiding overfitting because of the small amount of data in the training data set in the process of deep learning neural network training, the existing training set needs to be enhanced and expanded. The image sizes of high-resolution

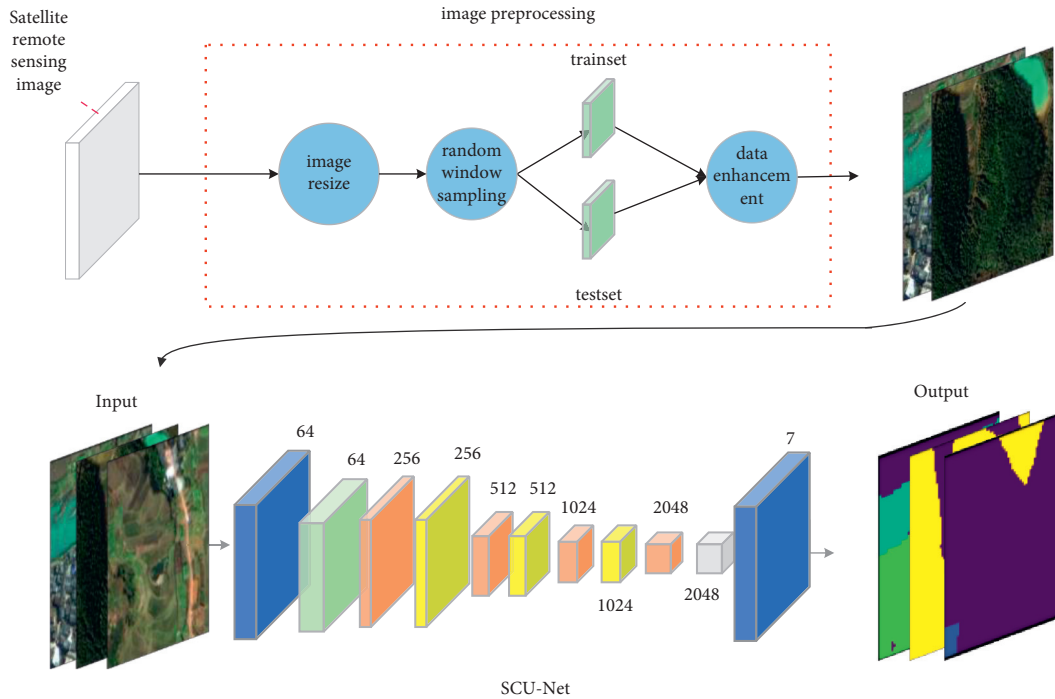


FIGURE 5: Remote sensing image segmentation flowchart (the convolutional layer is represented by blue blocks, the pooling layer is represented by green blocks, the superimposed convolutional layer is represented by orange blocks, the yellow block represents the CFWE, and the gray block represents the upsampling operation).

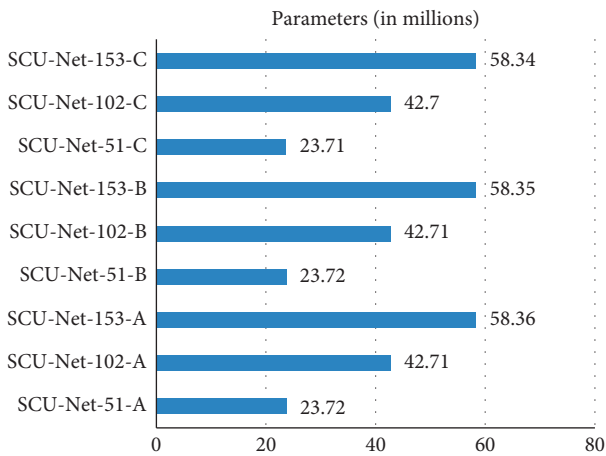


FIGURE 6: Comparison of SCU-net’s parameters.

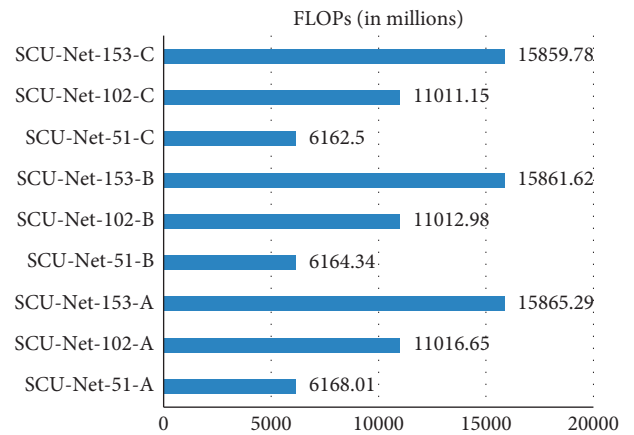


FIGURE 7: Comparison of floating-point operations (FLOPs).

remote sensing datasets are also inconsistent. Therefore, for a high-resolution remote sensing image, it needs to be cut into small images of uniform size as an experimental dataset. The pretreatment process is as follows:

- (1) Firstly, set the dataset image to the output image of the specified size as needed.
- (2) Secondly, perform random window sampling with a window size of 256×256 on the image, i.e., randomly generate sampling coordinates, and then obtain an image with a size of 256×256 under the coordinates.
- (3) Finally, data enhancement is performed on the segmented images, and image transformation

operations are performed, such as rotation, vertical flipping, horizontal flipping, blurring, corrosion, random gamma transformation, bilinear filtering, and adding noise at random.

After the above data preprocessing and data expansion, the amount of data in the training set has become 6 times that of the original dataset, which reduces the risk of network overfitting in a way.

To ensure the reliability of the performance comparison of different network models, the experiments in this paper are carried out under the same development platform and hardware environment. The experimental equipment system is Windows 10, CPU (central processing unit) is Intel i7

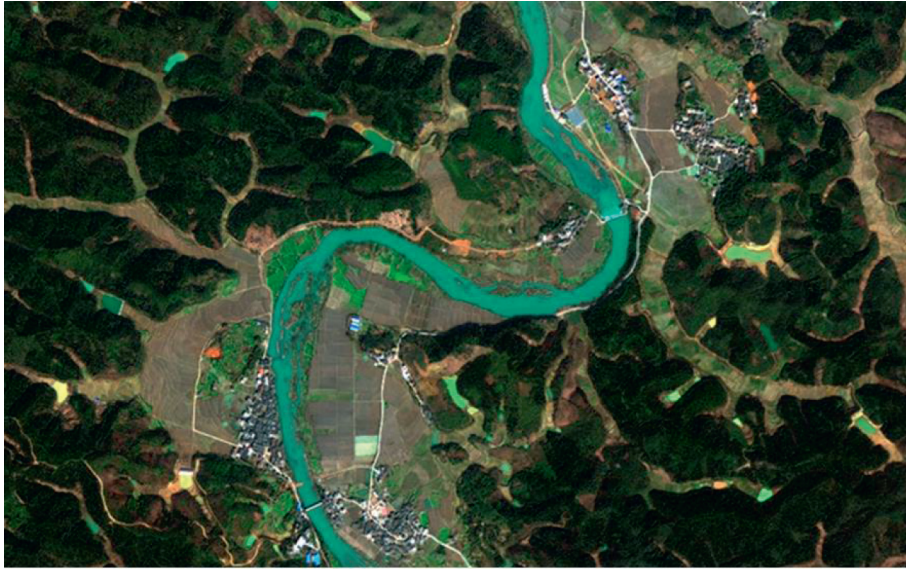


FIGURE 8: High-resolution remote sensing image.



FIGURE 9: Cases of satellite remote sensing images (the spatial resolution is 0.8 meters).

3.30 GHz, GPU (graphics processing unit) is GeForce GTX 1080Ti (11G), GPU environment is CUDNN 10.0 and CUDA 10.0, the development platform is PyCharm, the programming language is Python, and the framework is Pytorch. The batch size of the test set is 4, and the batch size of the training set is also 4.

3.3. Evaluation Criteria. There are a variety of evaluation criteria used in semantic segmentation models. The evaluation criteria adopted in this paper are those adopted by most semantic segmentation models, including frequency

weighted intersection (FWIU), pixel accuracy (PA), and mean intersection (MIOU) as performance indicators [4]. The evaluation criteria of frequency weighted intersection-over-union are defined as multiplying the frequency of each category as the weight with the intersection-over-union of each category and finally to sum. The evaluation criteria of pixel accuracy are defined as the ratio of correctly marked pixels to the total pixels. The evaluation criteria of mean intersection-over-union refer to the ratio of the intersection and union of the two sets of real and predicted values. Assume that the number of pixel categories is $K + 1$. The formulas of PA, MIOU, and FWIU are as follows:

$$\begin{aligned}
PA &= \frac{\sum_{i=0}^k P_{ii}}{\sum_{i=0}^k \sum_{j=0}^k P_{ij}}, \\
MIOU &= \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}}, \\
FWIU &= \frac{1}{\sum_{i=0}^k \sum_{j=0}^k P_{ij}} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}}.
\end{aligned} \tag{11}$$

Among them, P_{ii} represents the number of class i pixels that are correctly classified, and P_{ij} represents the number of class i pixels classified as class j .

In addition, the precision, recall, and F1-score indicators in image classification are used in the experiment to evaluate the prediction effect of a single category. TP is defined as the model correctly predicted class i . TN is defined as the negative sample predicted correctly by the model. Defining FP as a negative sample is incorrectly predicted by the model to be a positive sample. Define FN as the class i is predicted to be other classes. Then, the precision is defined as follows:

$$Precision = \frac{TP}{TP + FP}. \tag{12}$$

Recall is defined as follows:

$$Recall = \frac{TP}{TP + FN}. \tag{13}$$

F1-score is defined as the harmonic mean of precision and recall, which is as follows:

$$F_1 = \frac{1}{2} \left(\frac{1}{Precision + Recall} \right). \tag{14}$$

3.4. Experimental Results. In this paper, nine kinds of SCU-Nets are designed and tested on the dataset. The effect of depth and upsampling methods on the segmentation performance of SCU-Nets is studied. The experimental results are displayed in Table 1. The optimal experimental results are bolded.

According to the above table, the performance of SCU-Net using the C structure as the upsampling method is lower than the model using the other two structures. SCU-Net-102-A has the best comprehensive performance in the dataset. PA and MIOU are the highest, i.e., 91.53% and 55.84%, respectively. The MIOU of SCU-Net-102-A is the highest, which is 55.84%. It is 0.95% and 1.43% higher than SCU-Net-51-A and SCU-Net-153-A individually.

The results show that the depth of SCU-Net should be moderate, while scaling shallower or deeper may cause the performance of the network model to decrease. Meanwhile, SCU-Net-51-A is compared with other networks for studying the influence of parameters on network performance. Further comparisons are also conducted with the classic semantic segmentation neural networks DeepLab, U-Net, SegNet, FCN, Dilated, and the most recent networks FC-DenseNet and DFCN121. For the model obtained when

TABLE 1: The performance of different depth SCU-nets.

Model	PA (%)	MIOU (%)	FWIU (%)
SCU-net-51-A	91.30	54.99	85.95
SCU-net-102-A	91.53	55.84	85.83
SCU-net-153-A	91.58	54.51	85.98
SCU-net-51- B	91.16	54.56	85.62
SCU-net-102- B	91.43	54.76	85.69
SCU-net-153-B	91.10	54.54	85.16
SCU-net-51-C	91.25	53.74	85.49
SCU-net-102-C	91.06	54.37	85.29
SCU-net-153-C	91.15	54.24	85.45

the CFWE module in SCU-Net-102-A was deleted, a comparative experiment was carried out to verify the effectiveness of the CFWE module. The experimental comparison results are shown in Table 2. The optimal experimental results are bolded.

As can be seen from the above table, SCU-Net-102-A without a CFWE module is significantly lower than that of SCU-Net-102-A with CFWE module, and the parameters of the three CFWE modules are only 10,000. Hence, it can be inferred that the CFWE module has very few parameters and can greatly improve the performance of SCU-Net. DeepLabV3 uses and improves the ASPP module, which has a good effect on the GF-2 dataset, however, the performance is still lower than SCU-Net, and DeepLabV3 has many more model parameters than SCU-Net-102-A. FCN-8s achieved pixel classification by abandoning the fully connected layer of the traditional neural network and made great success on the GF-2 dataset with very few parameters. SegNet is a semantic segmentation network that improves VggNet-16 and is based on the FCN network. At the same time, it introduces the encoder-decoder structure, which has achieved good results on the GF-2 dataset, and the number of parameters is more than SCU-Net-51-A. U-Net was put forward to settle the matter of semantic segmentation of medical images. A U-shaped network structure was proposed to obtain location information and context information simultaneously. Although the parameters of U-Net are very few, it performs poorly on the GF-2 dataset. FC-DenseNet and DFCN121 introduced DenseNet into the model of semantic segmentation. At the same time, regardless of the number of other network parameters, SCU-Net performance is better than theirs. SCU-Net-51-A has low parameters and high performance. It shows that SCU-Net has better performance and lower model complexity and can be more targeted to complete the task of semantic segmentation of satellite remote sensing images.

To evaluate the application effect of the network in actual remote sensing images, use 2000 remote sensing images of 256×256 size as a test. Table 3 lists the predicted pixels corresponding to 7 classes in the test images in detail. The optimal experimental results are bolded. Class 0–6 in the dataset, respectively, represent the background (cultivated land), forest land, others, water area, buildings, road, and furrow. The best experimental results are shown in bold. Meanwhile, the precision, recall, and F1-score of SCU-Net-102-A and DFCN121C models are compared. The precision

TABLE 2: Performance of different depth SCU-nets.

Model	PA (%)	MIOU (%)	FWIU (%)	PRM (m)
SCU-net-102-A	91.53	55.84	85.83	42.71
NoSCU-net-102-A	91.36	53.77	85.52	42.70
DeepLabV3 [20]	87.87	49.70	81.80	58.00
FCN-8s [10]	90.55	51.16	84.61	14.70
SegNet [17]	91.17	50.60	85.35	28.40
U-net [16]	89.65	44.50	83.55	13.40
FC-DenseNet [11]	84.99	51.67	90.85	9.40
DFCN121 [13]	91.28	54.03	85.69	17.00
SCU-net-51-A	91.30	54.99	85.95	23.72

and recall retain two decimal places, and the F1-score retains three decimal places.

It can be seen from Table 3 that all kinds of F1 scores predicted by the SCU-Net-102-A model are higher than those predicted by the DFCN121C model, except the “others” class. In addition, the forest F-score predicted by the SCU-Net-102-A model is the highest, which is 0.952. The corresponding forest land predicted precision reached 97.10%. At the same time, it can be found that the prediction effect of others in Table 3 is not so ideal mainly because of the combination of classes during manual labeling. There are not many experimental data samples of bare land or wasteland, and the geographical characteristics are not very obvious. Hence, the prediction results will be affected to some extent.

3.5. Experiments Analysis. The experimental results show that the PA and MIOU of SCU-Net-102-A are the highest, which are 91.53% and 55.84%, respectively. After comprehensive consideration, SCU-Net-102-A is chosen to compare with other typical semantic segmentation networks and the more recent semantic segmentation networks. It is found that the comprehensive performance of the network is better than other networks. Judging from the experimental results, the network depth should be kept appropriate in the task of semantic segmentation of satellite remote sensing images. Too few network layers make it difficult to extract sufficient features, and too many network layers will cause vanishing gradient problem or gradient explosion problem. The problems of gradient dispersion and gradient explosion can be solved to some extent by batch normalization, and the degradation problem can be alleviated by jumping connection. The CDeConv upsampling module uses a point convolutional layer, which improves the feature extraction capability of the network on the basis of reducing the number of network parameters. In addition, the attention mechanism module CFWE module, which uses skip connections to settle the matter of network degradation to some extent, is introduced. Simultaneously, the CFWE module learns the channel weight coefficients using the attention mechanism, which enhance the feature extraction capability of the network and achieves a better image segmentation effect. SCU-Net-102-A and DFCN121C networks are selected for experiments to compare and predict the effects of different categories. The precision of the SCU-Net-102-A

model for forest land prediction is 97.10%, and the recall is 93.42%. The corresponding F1-scores reached 0.952, which is the best among the 7 classes, and these three values are higher than those of the DFCN121C model. At the same time, all kinds of F1 scores predicted by SCU-Net-102-A are higher than those predicted by DFCN121C, except for the “others” class. Combined with the experimental results, it can be proved that SCU-Net-102-A is better than DFCN121C in predicting each.

3.6. Visualization of Segmentation Results. The visualization diagram of the SCU-Net-A test output is shown in Figure 10. The first row is the input image of the test set, the second row is the ground truth, the third row is the image predicted by DFCN121C, and the fourth row is the image predicted by SCU-Net-102-A. In the experiment, two remote sensing images with a size of 1856×1856 were selected, and the SCU-Net-102-A and DFCN121C network models were used for prediction. The red rectangle indicates the prediction error, the yellow rectangle indicates the prediction effect on the water category, and the rose-red rectangle indicates the prediction effect on the road category.

On the whole, the prediction effect of SCU-Net-102-A is better than the image predicted by DFCN121C. In the first raw image, the waters are marked as bare ground because of human labeling errors. Both SCU-Net-A and DFCN121C can correctly predict waters, however, it is clear that SCU-Net-102-A understands the details better, as shown by the yellow boxes in the first column of Figure 10. As can be seen from the above figure, both SCU-Net-102-A and DFCN121C can predict unlabeled waters, however, the boundary information of SCU-Net-102-A is better than that of DFCN121C. It is indicated by the yellow box in the second column of Figure 10. Furthermore, it is experimentally found that DFCN121C predicts more misinformation than SCU-Net-102-A, as shown by the red boxes in Figure 10. In addition, there are unmarked roads in the original image in the second column, however, both SCU-Net-102-A and DFCN121C have learned this road information, and it can be seen from the figure that SCU-Net-102-A is better than DFCN121C as it learned more about the road category, as shown by the rose-red box in Figure 10. It shows that SCU-Net-102-A has a strong generalization ability in the semantic segmentation of high-resolution remote sensing images.

TABLE 3: Labeling results of SCU-net-102-A and DFCN121C for each class on the GF-2 dataset.

Model		Cultivated land	Forest land	Others	Water area	Building	Road
Predict pixels		22288550	92223235	199063	11547582	2982585	1830985
Predict pixels (TP)	SCU-net-102-A	15161346	89547666	113169	11183917	886270	1517840
	DFCN121C	14883805	89361908	117344	11255276	1999480	289870
Predict pixels (FP)	SCU-net-102-A	7127204	2675569	85894	363665	2096315	313145
	DFCN121C	7404745	2861327	81719	292306	983105	1541115
Precision	SCU-net-102-A	68.02%	97.10%	43.15%	96.85%	70.29%	82.90%
	DFCN121C	66.78%	96.90%	58.95%	97.47%	67.04%	84.17%
Recall	SCU-net-102-A	82.88%	93.41%	50.98%	92.15%	86.90%	72.87%
	DFCN121C	83.20%	93.16%	50.11%	90.71%	86.89%	70.87%
F1-score	SCU-net-102-A	0.747	0.952	0.467	0.944	0.777	0.776
	DFCN121C	0.741	0.949	0.541	0.939	0.757	0.769

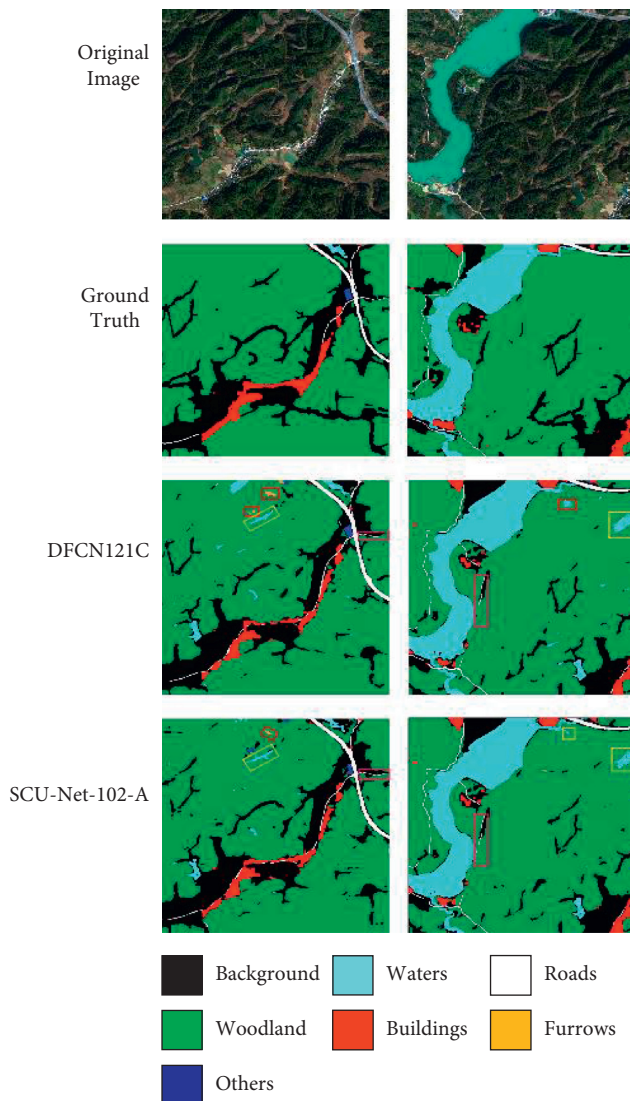


FIGURE 10: Prediction maps of the SCU-net-102-A and DFCN121C (the ground truth map is divided into 7 classes: 0 for background, 1 for woodland, 2 for bare land or wasteland, 3 for waters, 4 for buildings, 5 for roads, and 6 for furrows). The first row is the original high-resolution remote sensing image, the second row is the ground truth, the third row is the DFCN121C prediction map, and the fourth row is the SCU-net-102-A prediction map).

4. Conclusions

This paper proposes a new semantic segmentation model for high-resolution remote sensing images based on CNNs. A new upsampling module CDeConv is proposed in this network model, and the attention module CFWE is used to improve the feature extraction ability of the network. In the experiment of this paper, the GF-2 satellite remote sensing image of Chenzhou in 2016 is used, and the ground truth was labeled manually. Finally, there were 7 categories of classified objects in the dataset. By analyzing the experimental results, it is concluded that SCU-Net-102-A has the highest value, and its PA and MIOU are 91.66% and 55.61%, respectively. The precision of SCU-Net-102-A for woodland prediction is 97.10%, and the corresponding F1-score is 0.952, which is the highest F1-score among the seven categories. It can achieve accurate segmentation of complex targets in GF-2 image classification tasks. The model designed in this paper is designed to detect the changes in forest land and water area in Chenzhou, China, and to provide help for the protection and development of forest and water environment. It is hoped that more satellite remote sensing images of the same kind can be obtained in the future, and higher-quality segmentation datasets can be generated to further train SCU-Net to improve its versatility and segmentation performance.

Data Availability

The dataset can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the Natural Science Foundation of Hunan Province, China (2019JJ80105), the ChangSha Science and Technology Project (kq2004071), and the Hunan Graduate Student Innovation Project (CX20200882).

References

- [1] C. Yao, X. Luo, Y. Zhao, Z. Wei, and X. Chen, "A review on image classification of remote sensing using deep learning," in *Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, pp. 1947–1955, IEEE, Chengdu, China, December 2017.
- [2] Y. Liu, S. Piramanayagam, S. T. Monteiro, and E. Saber, "Semantic segmentation of multisensor remote sensing imagery with deep ConvNets and higher-order conditional random fields," *Journal of Applied Remote Sensing*, vol. 13, no. 1, Article ID 016501, 2019.
- [3] G. Baranoski, B. W. Kimmel, and P. Varsa, "Assessing the impact of porosity variations on the reflectance and transmittance of natural sands," *Journal of Applied Remote Sensing*, vol. 13, no. 2, Article ID 024522, 2019.
- [4] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 1097–1105, Montreal, Canada, December 2012.
- [5] W. Wang, Y. Yang, X. Wang, W. Z. Wang, and J. Li, "Development of convolutional neural network and its application in image classification: a survey," *Optical Engineering*, vol. 58, no. 4, Article ID 040901, 2019.
- [6] W. Wang, Y. Li, T. Zou, X. Wang, and Y. Luo, "A novel image classification approach via dense-mobilenet models," *Mobile Information Systems*, vol. 2020, Article ID 7602384, 2020.
- [7] W. Wang, H. Liu, J. Li, H. Nie, and X. Wang, "Using CFW-net deep learning models for X-ray images to detect COVID-19 patients," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, pp. 199–207, 2020.
- [8] Y. Xue, Y. Wang, J. Liang, and A. Slowik, "A self-adaptive mutation neural architecture search algorithm based on blocks," *IEEE Computational Intelligence Magazine*, vol. 16, no. 3, pp. 67–78, 2021.
- [9] Y. Xue, P. Jiang, F. Neri, and J. Liang, "A multi-objective evolutionary approach based on graph-in-graph for neural architecture search of convolutional neural networks," *International Journal of Neural Systems*, vol. 31, no. 9, Article ID 2150035, 2021.
- [10] 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.
- [11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [12] W. Wang, C. Tang, X. Wang, Y. Luo, Y. Hu, and J. Li, "Image object recognition via deep feature-based adaptive joint sparse representation," *Computational Intelligence and Neuroscience*, vol. 2019, Article ID 8258275, 9 pages, 2019.
- [13] Y. LéCun, 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.
- [14] K. Simonyan and A. Zisserman, "Very deep convolutional net-works for large-scale image recognition," arXiv preprint arXiv: 1409.1556v6, 2014.
- [15] C. Szegedy, W. Liu, and Y. Jia, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, 2015.
- [16] K. He, X. Zhang, and S. Ren, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, 2016.
- [17] J. Simon, M. Drozdal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1175–1183, IEEE, Honolulu, HI, USA, July 2016.
- [18] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2261–2269, Honolulu, HI, USA, July 2017.
- [19] W. Wang, Y. J. Yang, J. Li, Y. Hu, and W. Wang, "Woodland Labeling in Chenzhou, China via deep learning approach," *International Journal of Computational Intelligence Systems*, vol. 13, pp. 1393–1403, 2020.
- [20] A. Hamida, A. Benoit, P. Lambert, and L. Klein, "Deep learning for semantic segmentation of remote sensing images with rich spectral content," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pp. 2569–2572, Chennai, India, April 2017.
- [21] W. Wang, Y. Jiang, Y. Luo, J. Li, and T. Zhang, "An advanced deep residual dense network (DRDN) approach for image super-resolution," *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 1592–1601, 2019.
- [22] 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*, p. 3, 2015.
- [23] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 2481–2495, 2017.
- [24] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected CRFs," *Computer Science*, vol. 4, pp. 357–361, 2014.
- [25] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.
- [26] L. C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," arXiv:1706.05587, 2017.
- [27] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image Segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 833–851, Springer, Munich, Germany, September 2018.
- [28] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800–1807, IEEE, Honolulu, HI, USA, July 2017.
- [29] P. Mitra, B. Uma Shankar, and S. K. Pal, "Segmentation of multispectral remote sensing images using active support vector machines," *Pattern Recognition Letters*, vol. 25, no. 9, pp. 1067–1074, 2004.
- [30] G. Bilgin, S. Erturk, and T. Yildirim, "Segmentation of hyperspectral images via subtractive clustering and cluster validation using one-class support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 8, pp. 2936–2944, 2011.

Research Article

Research on Urban National Sports Fitness Demand Prediction Method Based on Ant Colony Algorithm

Wei Yue ¹ and Peng Dai ²

¹College Physical Education Department, Qingdao Huanghai University, Qingdao, Shandong 266427, China

²Dean's Office, Qingdao Binhai University, Qingdao, Shandong 266427, China

Correspondence should be addressed to Wei Yue; yuew01@qdhhc.edu.cn

Received 12 February 2022; Revised 17 March 2022; Accepted 24 March 2022; Published 7 April 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Wei Yue and Peng Dai. 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 competitive sports, the enthusiasm of the public to participate in sports has gradually increased. At present, almost all streets in the city have their own fitness places, which provide a lot of help for public fitness. However, the existing fitness venues are obviously insufficient, the venues are limited, relatively single, and the open-space area is insufficient, which cannot meet the needs of mass sports fitness. Based on this, this paper studies and analyzes the prediction of urban national sports fitness demand based on the ant colony algorithm. First, this paper analyzes the National Fitness Situation and the related research on demand forecasting and puts forward the use of the ant colony algorithm to realize demand forecasting. This paper expounds on the research methods and algorithms commonly used in demand forecasting. The ant colony algorithm is used to improve the fuzzy analysis. The urban national sports fitness demand is divided into six secondary indicators, and different tertiary indicators are divided under each secondary indicator. Through simulation analysis, it is confirmed that the improved algorithm proposed in this paper converges faster and finds the best path most. At the same time, the weight of the urban national sports fitness demand index is calculated.

1. Introduction

The ability to deal with public health emergencies is an urgent need to maintain national security and social stability. Timely and stable responses to public health emergencies are related to national security and development and the overall stability of the economy and society. If it is not handled properly and controlled effectively, not only the people's living standards and quality will be greatly affected but also people's panic will be caused, and the whole society will pay a heavy price. Improving the ability to deal with public health emergencies is the internal requirement to improve the national governance ability. Preventing and responding to public health emergencies is a complex and highly interrelated systematic project, which is a major test of the national governance system and governance ability.

With the improvement of the national economic level, the public's health awareness has also been significantly

enhanced. At present, people are eager to improve their physical quality by participating in physical exercise. Therefore, the social demand for physical fitness is also accumulating [1]. Moreover, with the development of competitive sports, people began to pay attention to the development of sports and have a certain demand for sports products and venues. At present, in the research on sports fitness demand, the early research mostly emphasized the necessity and inevitability of health demand and also made it clear that sports health demand can bring a lot of economic value to the society [2]. In the relevant literature on demand content, in addition to analyzing the inevitability of sports health demand, the research on education also began to become a hot spot. With the acceleration of the aging process, the universal public health mode has been popularized [3]. Some scholars have also analyzed the differences in mass sports fitness needs and believe that there are great differences in the development of competitive sports

between urban and rural areas, and the suburbs have great advantages in organizing events and activities. In these studies, qualitative analysis is mostly used, and quantitative questionnaire survey is mainly used, which is lack of predictability [4]. Based on this, this paper studies the prediction method of urban national health sports demand based on the ant colony algorithm.

This paper mainly studies and analyzes the urban national sports fitness demand prediction based on the ant colony algorithm, which is mainly divided into four chapters. Chapter 1 focuses on the research background and overall framework of this paper. The second part focuses on the research methods of demand forecasting at home and abroad and the application of the ant algorithm. The third chapter constructs the urban national sports fitness demand prediction, uses the fuzzy algorithm to layer the fitness demand prediction and calculate the weight, and improves the ant colony algorithm on the fuzzy analysis to realize the rapid calculation of the weight. The fourth chapter carries out simulation analysis, verifies the effectiveness of the algorithm proposed in this paper through the indicators such as convergence speed, convergence result, and error, and calculates the weight of the prediction index of urban national sports fitness demand.

Compared with the previous research results, the innovation of this paper lies in the analysis of research methods. Most of the previous studies used the literature method or questionnaire survey method, which is lack of accuracy and heavy workload. This paper proposes to use the ant colony algorithm to realize demand prediction analysis. Through the analysis of previous literature, the national sports fitness demand is divided into multiple secondary indicators, and the secondary indicators are divided into different tertiary indicators. Ant colony algorithm is used to improve the fuzzy algorithm and calculate the weight of tertiary indicators, which can compare and analyze the importance of different indicators. Then, it analyzes the concentrated performance of national fitness demand, which has a certain reference significance for improving the construction of stadiums and sports culture development.

2. Related Work

In recent years, there are many researches on demand forecasting, and some scientific research achievements have been made [5]. In the research and analysis, Alireza et al. used the ant colony algorithm to analyze the highway network of Kerman province, the largest province in Iran, which can identify the existing bottlenecks and predict the potential bottlenecks in the future [6]. Fan y et al. proposed a new model with unservice requirements by relaxing the requirements of meeting constraints, designed an algorithm based on distributed ant colony optimization (ACO), and made some specific modifications to it, using the average cost-saving percentage of each bicycle as a measure to evaluate the performance of our method in reducing costs [7]. Prasad et al. designed a method combining multistage multivariable empirical mode decomposition with ant colony optimization and random forest (i.e., MEMD-ACO-RF)

to predict monthly solar radiation (RN). After calculating multivariable IMF, the ant colony optimization (ACO) algorithm was used to determine the best feature based on IMF for model development by combining the historical lag data at $(T - 1)$. The RF model is applied to the selected IMF to predict monthly RN [8]. In price reservation, Meng et al. proposed a parametric pricing method to formulate the pricing variables representing the pricing factors, calculated the pricing variables according to VMM and the regression relationship between the pricing variables and the price, and proposed a support vector regression (SVR)-integrated price prediction based on ant colony optimization algorithm (ACO) [9]. In climate research, Chaudhuri S et al. introduced swarm intelligence in the form of ant colony optimization (ACO) technology to calculate pheromone deposition on the path of the tropical cyclone and then used a neural network to predict the maximum sustained wind speed of cyclone over the Bay of Bengal (NiO) in the northern Indian Ocean [10]. In the research and analysis, Wan young et al. took baseball, bicycle, golf, and hiking club members as the research objects and analyzed the consumption tendency of sporting goods of activity participants from multiple dimensions [11]. In their research, Zhang y and others proposed a neural network sports load prediction model combined with the ant colony algorithm. The global search ability of the ant algorithm is used to determine the initial weight of the neural network, and the weight is further adjusted on the basis of neural network gradient descent to find the global best advantage. In the experiment, a three-layer BP neural network is used. It absorbs the advantages of the ant colony algorithm and neural network and has obvious advantages [12].

To sum up, it can be seen that there are many research methods related to demand forecasting, including cluster analysis, ant colony algorithm, and vector machine. In terms of research objects, in addition to theoretical research, demand forecasting research has a wide range of applications, such as logistics model, electricity, and transportation. However, there are few studies on sports, and there are few studies on the prediction of mass sports fitness demand. The prediction of sports demand is only involved and not in-depth. On the other hand, the existing prediction of physical exercise demand is generally based on a questionnaire survey, and the combination with an algorithm is generally related to site selection. Few studies can combine ant colony algorithms to construct sports health demand prediction.

3. Methodology

3.1. Neural Network Algorithm Based on Ant Colony Algorithm. The analytic hierarchy process is widely used in evaluation model, but this algorithm has a slow convergence speed and is easy to fall into a local minimum. This paper studies the use of the ant colony algorithm to improve the urban national sports fitness demand analysis. Ant algorithm simulates the ant's lost behavior and the ability to find the best nearest road, makes full use of the mechanism of selection, update, and coordination, and finds the optimal solution through the information exchange between

individuals. Therefore, the ant colony algorithm is the algorithm that is most easy to find the better solution [13]. This paper uses the ability of the ant colony algorithm to add pheromone to each weight and uses the mean square error to simulate finding the optimal weight combination. Assuming that the number of ants is m , the obtained component is regarded as a vertex, and each vertex is a component. There are ki connecting lines between the i and $i + 1$ vertices, which means that the value of the component is within this range [14]. Assuming that the pheromone on the j route at t time is $t_{ij}(t)$, the ant starts from this vertex, reaches the next vertex according to the strategy, and finally reaches the n vertex. The path that the ant passes through is a scheme. Assuming that the probability of ant transfer is expressed by P , the formula is

$$P_{ij}^k(t) = \begin{cases} \frac{t_{ij}^\alpha \eta_{ij}^\beta(t)}{\sum_{0^r \in \text{allowed}_k} t_{ij}^\alpha \eta_{ij}^\beta(t)}, & j \in \text{allowed}_k, \end{cases} \quad (1)$$

where η belongs to heuristic information and represents visibility, t represents the strength of residual pheromone, and P represents the probability of ants choosing this path.

The path selected by the ant in the next step can be expressed as $\text{allowed}_k = \{n - \text{tabu}_k\}$. After a period of time, a cycle is completed, and the pheromone of the path will change. The formula is adjusted to

$$t_{ij}(t+1) = (1-p)t_{ij}(t) + p\Delta t_{ij}(t, t+1),$$

$$\Delta t_{ij}(t, t+1) = \sum_{k=1}^m \Delta t_{ij}^k, \quad (t, t+1), \quad (2)$$

where p represents the pheromone volatilization coefficient, which can reduce the continuous accumulation. The value range is 0~1, and Δt represents the pheromone left by the ant after selecting a path. In this formula, due to different algorithms, it is divided into three types, namely,

$$\Delta t_{ij}^k(t, t+1) = \begin{cases} Q, & \text{cycle passes through I and J,} \\ 0, & \text{else,} \end{cases}$$

$$\Delta t_{ij}^k(t, t+1) = \begin{cases} \frac{Q}{d_{ij}}, & \text{cycle passes through I and J,} \\ 0, & \text{else,} \end{cases} \quad (3)$$

$$\Delta t_{ij}^k(t, t+1) = \begin{cases} \frac{Q}{L_k}, & \text{cycle passes through I and J.} \\ 0, & \text{else.} \end{cases}$$

The first two algorithms release pheromones during construction, and the third one releases pheromones after completion [15]. The longest path traveled by ants is used as the overall information, and the pheromone strength is constant, which will affect the convergence speed.

When the ant colony algorithm is used for demand forecasting, it is mainly used to complete the premise work to optimize the weights and thresholds of the neural network, narrow the range, and then accelerate the convergence in combination with the improvement of the neural network [16], as shown in Figure 1. Before the implementation of the ant colony algorithm, each weight variable of the neural network is segmented. The ant passes through each weight vector, selects a value in the definition domain, completes the selection of weight variables, constructs a group of weights, inputs the sample data, and updates the pheromone after calculating the error. Using the ant algorithm can improve performance and deal with a wide range of data sets [17]. In order to better deal with the data, we need to deal with the prior knowledge of the problem. In order to better explore the path, the principle of random proportion is

adopted. The global update is carried out on the optimal ant path. After each cycle, only the path traveled by the ant is enhanced. In this way, under the volatilization mechanism, other paths are continuously reduced, the difference between the optimal path and other paths is completed, the optimal path is reached more quickly, and the efficiency of the algorithm is quickly improved [18]. Specifically, when selecting the path, the random proportion principle is used, and the ants at the node use the rules to select the point. The formula is

$$s = \begin{cases} \arg \max_{u \in \text{allowed}_k} \{[t(i, u)]^a [\eta(i, u)]^\beta\}, & q \leq q_0, \\ s, & \text{else,} \end{cases} \quad (4)$$

where i represents the node, q represents the randomly generated value, ranging from 0 to 1, and S represents the proportion rule. This random rule can use the prior knowledge about the problem to limit the search to the range of the optimal solution and can use the proportional rule to know the relative importance. When ants choose the path, they give priority to the path with the shortest path and the

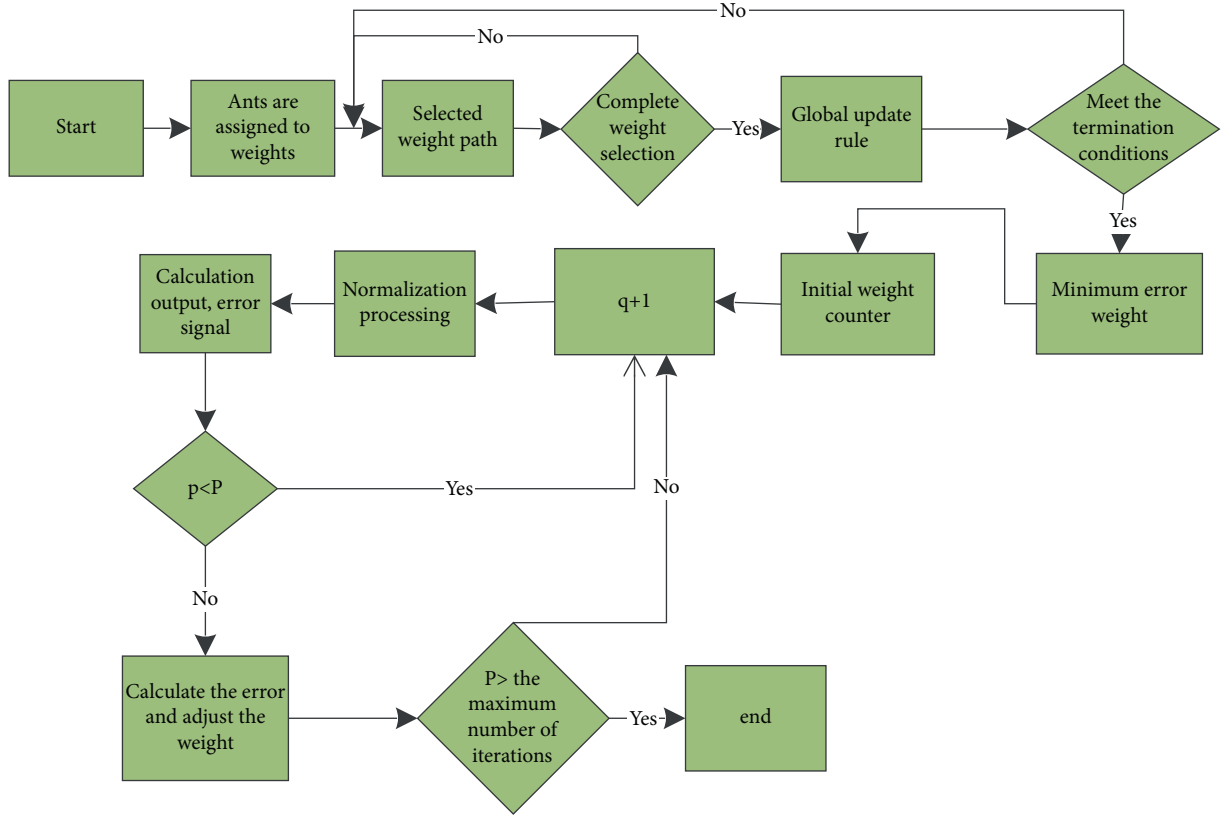


FIGURE 1: Network optimization flow chart of ant colony algorithm.

most pheromones. It is assumed that the preset parameters are expressed by a and b to describe the relationship between heuristic information and pheromone concentration. After each cycle is completed, the optimal information degree is enhanced, and the global optimal path forms a weight

sequence [19]. The minimum error of the sequence is calculated. Assuming that the volatilization parameter is expressed by P , the pheromone update formula can be expressed as

$$t(i, s) = (1 - p)t(i, s) + p\Delta t(i, s),$$

$$\Delta t(i, s) = \begin{cases} \frac{Q}{E}, & \text{path belongs to the optimal path,} \\ 0, & \text{else,} \end{cases} \quad (5)$$

where, if (i, s) belongs to the global optimal path, $\Delta t(i, s) = Q/E$; otherwise, $\Delta t(i, s) = 0$.

In the application of a neural network based on the ant colony algorithm, the weights are initialized and divided first. According to the random ratio rules, the ants transfer and mark the passing points. When the ownership value is taken, a cycle is completed, and the recorded points form a group of weights to obtain the data and error. The pheromone is updated according to the global update strategy until the maximum evolution algebra is satisfied, and the adaptive learning rate improvement algorithm is used for further calculation. Through these calculations, the

convergence speed of the algorithm can be quickly improved. However, due to the characteristics of the ant colony algorithm itself, the search speed is still slow, so it needs to be further improved to avoid premature [20]. After each cycle, only one ant pheromone is set to update to find the optimal solution for the current cycle. In order to avoid stagnation, all elements are limited to a fixed interval. Pheromone trajectory rules can be expressed as

$$t_{ij}(t+1) = \rho t_{ij}(t) + \Delta t^{\text{best}_{ij}}, \quad (6)$$

where t_{\max} represents the trajectory and f represents the iterative optimal solution. The pheromone evaluation mechanism is further updated, and the formula is expressed as

$$t_{ij}^*(t) = t_{ij}(t) + \delta(t_{\max}(t) - t_{ij}(t)), \quad (7)$$

where δ represents the retention of previous pheromones, ranging from 0 to 1. When $\delta = 1$, it indicates trajectory initialization, and when $\delta = 0$, it indicates a complete shutdown mechanism. Under this mechanism, the pheromone distribution at stagnation can be obtained.

In the incremental allocation of information, different paths get the same enhancement, and the search results are different in different sections. For better paths, larger pheromones need to be allocated, and for worse paths, smaller pheromones need to be allocated. Assuming that the total number of occurrences of the path in the search cycle is k , there is

$$\Delta t_{ij} = \frac{\sqrt[k]{k}}{f(s^{\text{best}})}, \quad (8)$$

where $f(s^{\text{best}})$ represents the optimal path length. In order to speed up the process and promote ants to gather in the optimal path as soon as possible, weakening the worst solution is considered. On the nonoptimal and worst paths, the pheromone is adjusted according to the following formula:

$$t_{ij}(t+1) = \rho t_{ij}(t) - \Delta t^{\text{worst}_{ij}}, \quad (9)$$

where ρ is the parameter, and the value is 0.5.

3.2. Fitness Demand Prediction Model. In the new era, the evaluation index system of urban national sports fitness needs to meet the requirements of scientificity, operability, and reliability. The corresponding index values of many indicators need to be able to be analyzed quantitatively. In the demand for sports health, it is necessary to put forward some data that are easy to obtain but cannot obtain specific indicators and reasonably match some indicators. In the evaluation of national sports fitness needs, we hope to cover most cities. However, due to the great differences in economic levels, many indicators should take the average value. In the selection of indicators, we should combine the professionalism and practicality of sports and improve adaptability at the same time.

Combined with the construction of physical fitness, the construction level of physical fitness is the objective subject, and there are six secondary indicators, namely fitness demand, entertainment demand, talent demand, product demand, social demand, and venue demand. Each indicator corresponds to the tertiary indicators, as shown in Figure 2. Fitness needs include exercise times and practice, need for venues, and need for environment and clubs. Entertainment needs are divided into the hope that projects will be valued, need more practice, need to improve their level, watch sports projects, and participate in sports teams. Talent needs are divided into high-level athletes' joint participation, sports instructors, managers, sports talents, and sports coaches. The demand for sports products is divided into purchasing

sports products, increasing product consumption, better shopping, and liking sports clothes. The social demand is divided into enhancing the feelings with friends through sports, learning to communicate, increasing the scope of making friends, and requiring the joint participation of friends. The venue demand is divided into more sports places, reducing unnecessary places, open places, more sound surrounding settings, and more space for movement.

Fuzzy comprehensive evaluation method is an improvement of the previous mandatory evaluation. This method divides the evaluation area into different types and makes weighted analysis on each index, which belongs to the evaluation method of the combination of qualitative analysis and quantitative analysis. The fuzzy evaluation method is more effective for complex systems. When applied, the definition model is inputted, the weight coefficient of each index by the analytic hierarchy process is calculated, and the weight of each index is expressed by w_i . The weight of secondary indicators is represented by w_{ij} , and the weight of three-pole indicators is represented by w_{ijk} . The indexes of sports demand are dimensionless to ensure that different indexes can be compared, and the index analysis model is obtained. All indicators are positive, so the fuzzy membership function is used to standardize each indicator. The function is

$$\Phi(e_{ij}) = \begin{cases} 1, & e_{ij} = M_j, \\ \frac{e_{ij} - m_j}{M_j - m_j}, & m_j < e_{ij} < M_j, \\ 0, & e_{ij} = m_j, \end{cases} \quad (10)$$

where e_{ij} represents the specific attribute value of the regional indicator, M_j represents the maximum value of the indicator, and m_j represents the minimum value of the indicator. This quantitative method can ensure the comparability between different indicators. The larger the value, the closer the actual value is to the maximum value. The difference between 1 and 1 is the difference in the maximum demand. The following formula is used to calculate the sports hypothesis:

$$Y = f(i) = \sum_{j=1}^{48} \Phi(e_{ij})w_j, \quad (11)$$

where Φ represents the fuzzy membership function value of the index.

4. Result Analysis and Discussion

4.1. Simulation Analysis of Ant Colony Algorithm. TSPLIB data are selected, the experimental parameter α is 1, β is 3, ρ is 0.95, C is 0.01, P_{best} is 0.05, and the number of iterations is 3000 test 10 times each time and calculates the average value. The calculation results are shown in Figure 3. It can be seen from the data in the figure that the error between the algorithm proposed in this paper and the optimal solution is

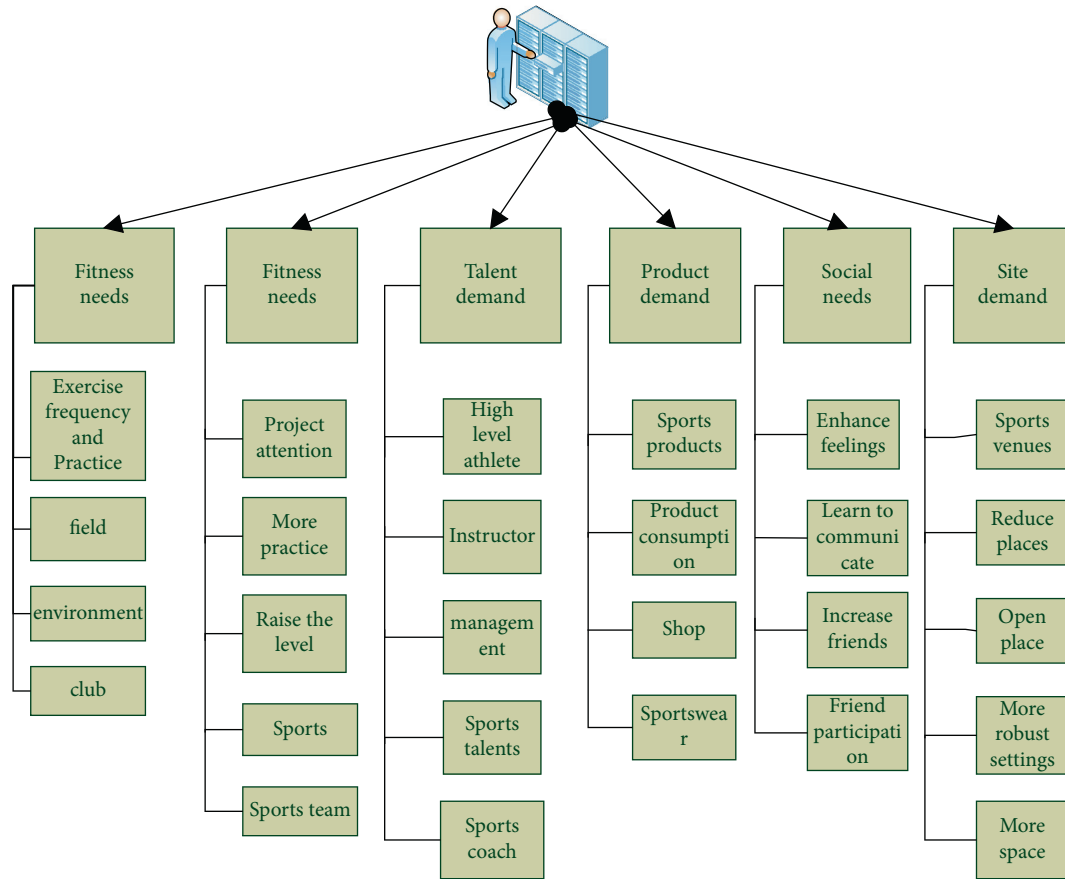


FIGURE 2: Sports fitness demand index system.

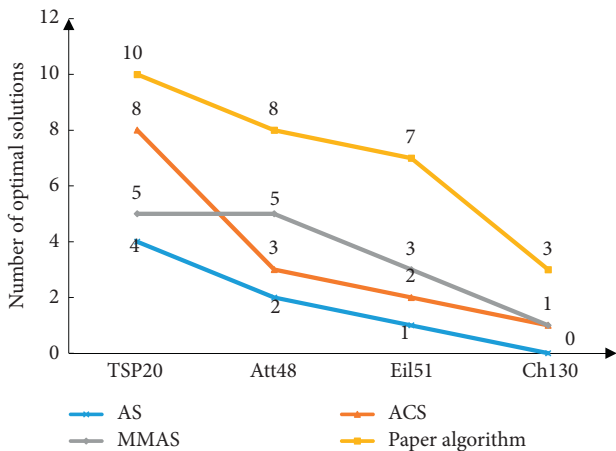


FIGURE 3: Performance comparisons of 4 algorithms.

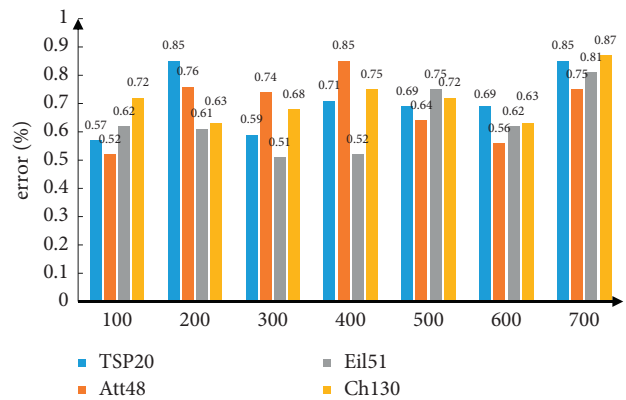


FIGURE 4: Error analysis.

the smallest in different examples, and the number of times to find the optimal solution is also the largest.

The optimal solution data published by TSPLIB are analyzed, and the optimal solution data are found by the algorithm in this paper and calculated the error, as shown in Figure 4. It can be seen from the figure that the error is less than 1%, which proves the effectiveness of the algorithm.

The robustness of the algorithm is analyzed in MATLAB 7. When the training error result reaches 0.1, it is considered

that the training is correct; otherwise, the training is continued. The weight variables are changed and modified in order to get better results. This paper tests the algorithm and iterates the loop. The convergence images are compared and analyzed, the error and convergence time are analyzed, and the internal initial weight is improved. The comparison results of internal improved convergence are shown in Figure 5. It can be seen from the data in the figure that the algorithm proposed in this paper can greatly improve the convergence.

The learning rate is set as 0.9 and 0.0001 as too small values and analyzes the correction results of the convergence

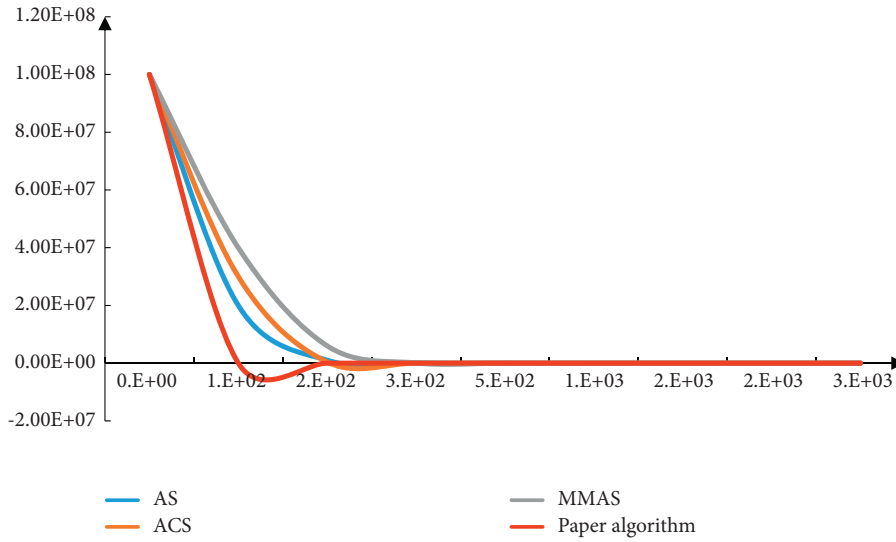


FIGURE 5: Convergence analysis of the algorithm.

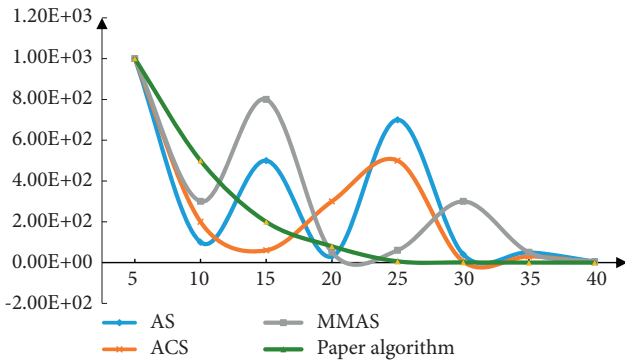


FIGURE 6: Correction result of excessive learning rate.

effect, as shown in Figure 6. From the changes in the figure, it can be seen that the algorithm proposed in this paper has an obvious convergence effect.

The analysis of the convergence time and results of the algorithm is shown in Figure 7. From the data in the figure, it can be seen that the convergence speed of the algorithm proposed in this paper is significantly improved. The scenic spot algorithm needs 1000 iterations to achieve the required accuracy. The improved algorithm can save 2/3 of the time and have less error.

4.2. Forecast and Analysis of Fitness Demand. The prediction indexes of urban national fitness demand are screened and revised, the optimal weight combination is studied and found, an adaptive algorithm is learned to avoid premature convergence, the performance of the ant colony algorithm is improved, and the national fitness demand is predicted and analyzed, which is divided into two-level indexes and three-level indexes. By comparing the indicators, a judgment matrix is built, the weight calculation is completed, and the final weight calculation result is obtained. In the process of index data processing, because the index itself has qualitative and quantitative analysis, it is necessary to

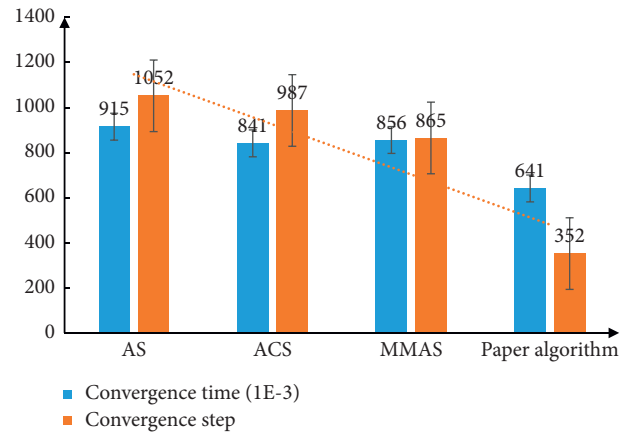


FIGURE 7: Convergence time and results of the algorithm.

process the data. All original indicators are quantified to ensure comparability between different indicators. The weight calculation is completed by MATLAB software, and the weight coefficient of each index is obtained through a consistency test.

Among the secondary indicators, the weight coefficient of fitness demand is 0.1, entertainment demand is 0.1, talent demand is 0.2, product demand is 0.3, social demand is 0.1, and venue demand is 0.2. Each indicator corresponds to three-level indicators. Under the demand of fitness, the exercise weight coefficient of the three-level indicators is 0.3. The required site weight coefficient is 0.3, the required environment weight coefficient is 0.3, and the club weight coefficient is 0.1. Under the secondary indicators of entertainment demand, the tertiary indicators hope that the project will be paid attention to. Under the talent demand index, the weight coefficient of high-level athletes participating in the three-level index is 0.2, the weight coefficient of sports instructors is 0.2, the weight coefficient of managers is 0.3, the weight coefficient of sports talents is 0.2, and the weight coefficient of sports coaches is 0.1. Under the secondary index of sports products, the weight coefficient of the

tertiary index for purchasing sports products is 0.2, the weight coefficient of increased product consumption is 0.2, the weight coefficient of better stores is 0.2, and the weight coefficient of favorite sports clothing stores is 0.4. Under the social needs of the secondary indicators, the weight coefficient of the tertiary indicators is 0.3, the weight coefficient of learning to communicate is 0.2, the weight coefficient of increasing the range of making friends is 0.2, and the weight coefficient of requiring the joint participation of friends is 0.3. Under the site demand of secondary indicators, the weight coefficient of more stadiums in tertiary indicators is 0.3, the weight coefficient of reducing unnecessary places is 0.1, the weight coefficient of open places is 0.2, the weight coefficient of surrounding settings is 0.3, and the weight coefficient of larger sports space is 0.1.

5. Conclusion

This paper studies the demand prediction of urban national sports fitness based on the ant colony algorithm and puts forward the shortcomings of the fuzzy analytic hierarchy process. The research adopts an ant colony algorithm to improve, increases pheromone for each group of weight, updates information with mean square error, and finds the optimal weight combination. And learn from the adaptive algorithm to avoid premature convergence, improve the performance of the ant colony algorithm, predict and analyze the demand for national fitness sports, and divide it into two-level indicators and three-level indicators. Through analysis, it is proved that the algorithm proposed in this paper has good convergence, effectiveness, and error performance, and the superiority of the algorithm is proved. It should be pointed out that there are many demand forecasting algorithms. This paper focuses on the analysis of the ant colony algorithm. At present, there is little research on the application of the algorithm in sports demand, which needs to be improved in combination with the specific problems in parameter selection. However, in the simulation analysis of this paper, the test is carried out in an ideal state. Affected by economic and environmental conditions, many data may not be continuous and complete. Therefore, the analysis of the algorithm needs to be further strengthened in future research.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] K. Foster, J. Stoeckle, A. Silverio et al., "Attitudes surrounding a community-based fitness intervention at an urban FQHC," *Family Medicine*, vol. 51, no. 7, pp. 598–602, 2019.
- [2] A. Fry and D. Putrino, "Exercise decreases cardiovascular risk factors: now what?" *International Journal of Cardiology*, vol. 254, no. 1, pp. 340–341, 2018.
- [3] B. Sekendiz, "Fit for kids," *ACSM's Health & Fitness Journal*, vol. 22, no. 3, pp. 33–36, 2018.
- [4] H. P. Sit, J. J. Yu, and W. Y. Huang, "Results from Hong Kong's 2018 report card on physical activity for children and youth," *Journal of Exercise Science & Fitness*, vol. 18, no. 3, pp. 1–17, 2020.
- [5] A. S. Ha, C. Lonsdale, D. R. Lubans, and J. Y. Y. Ng, "Increasing students' physical activity during school physical education: rationale and protocol for the SELF-FIT cluster randomized controlled trial," *BMC Public Health*, vol. 18, no. 1, p. 11, 2018.
- [6] M. Alireza and A. Amir, "A heuristic method to determine traffic bottlenecks based on ant colony: a case study of Iran," *Case Studies on Transport Policy*, vol. 6, no. 4, pp. 716–721, 2018.
- [7] Y. Fan, G. Wang, X. Lu, and G. Wang, "Distributed forecasting and ant colony optimization for the bike-sharing rebalancing problem with unserved demands," *PloS one*, vol. 14, no. 12, Article ID e0226204, 2019.
- [8] R. Prasad, M. Ali, P. Kwan, and H. Khan, "Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation," *Applied Energy*, vol. 236, pp. 778–792, 2019.
- [9] Q. N. Meng and X. Xu, "Price forecasting using an ACO-based support vector regression ensemble in cloud manufacturing," *Computers & Industrial Engineering*, vol. 125, no. 10, pp. 171–177, 2018.
- [10] S. Chaudhuri, D. Basu, D. Das, S. Goswami, and S. Varshney, "Swarm intelligence and neural nets in forecasting the maximum sustained wind speed along the track of tropical cyclones over Bay of Bengal," *Natural Hazards*, vol. 87, no. 3, pp. 1413–1433, 2017.
- [11] W. Lee and S. Au, "Study on the relationships between perceptions and patterns of consumption in sports subgroups," *The Korean Journal of Physical Education*, vol. 56, no. 2, pp. 63–76, 2017.
- [12] Y. Zhang, P. Zhang, and X. Sun, "Sports load prediction based on the improved neural network through ant colony algorithm," *CeCa*, vol. 42, no. 3, pp. 1073–1076, 2017.
- [13] N.-a. Mokhtari and V. Ghezavati, "Integration of efficient multi-objective ant-colony and a heuristic method to solve a novel multi-objective mixed load school bus routing model," *Applied Soft Computing*, vol. 68, pp. 92–109, 2018.
- [14] R. Goel and R. Maini, "A hybrid of Ant Colony and firefly algorithms (HAFA) for solving vehicle routing problems," *Journal of Computational Science*, vol. 25, no. 5, pp. 28–37, 2018.
- [15] Y. Qian, F. Hou, J. Fan, Q. Lu, X. Fan, and G. Zhang, "Design of a fan-out panel-level SiC MOSFET power module using ant colony optimization-back propagation neural network," *IEEE Transactions on Electron Devices*, vol. 68, no. 99, pp. 1–8, 2021.
- [16] M. Vafaei, A. Khademzadeh, and M. A. Pourmina, "A new QoS adaptive multi-path routing for video streaming in urban VANETs integrating ant colony optimization algorithm and fuzzy logic," *Wireless Personal Communications*, vol. 118, no. 6, pp. 1–34, 2021.
- [17] B. Ikhlef, C. Rahmoune, B. Toufik, and D. Benazzouz, "Gearboxes fault detection under operation varying condition based on MODWPT, Ant colony optimization algorithm and Random Forest classifier," *Advances in Mechanical Engineering*, vol. 13, no. 8, pp. 1–30, 2021.
- [18] P. Stodola, K. Michenka, J. Nohel, and M. Rybanský, "Hybrid algorithm based on ant colony optimization and simulated

annealing applied to the dynamic traveling salesman problem,” *Entropy*, vol. 22, no. 8, p. 884, 2020.

- [19] F. Wu, “Contactless distribution path optimization based on improved ant colony algorithm,” *Mathematical Problems in Engineering*, vol. 2021, no. 7, pp. 1–11, Article ID 5517778, 2021.
- [20] L. Wan and C. Du, “An approach to evaluation of environmental benefits for ecological mining areas based on ant Colony algorithm,” *Earth Science India*, vol. 14, no. 2, pp. 797–808, 2021.

Research Article

3D Virtual Modeling Realizations of Building Construction Scenes via Deep Learning Technique

Weihong Li 

Xi'an University of Finance and Economics, Xi'an, Shanxi 710100, China

Correspondence should be addressed to Weihong Li; liweihong@xaufe.edu.cn

Received 8 January 2022; Revised 4 February 2022; Accepted 17 February 2022; Published 31 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Weihong 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.

The architectural drawings of traditional building constructions generally require some design knowledge of the architectural plan to be understood. With the continuous development of the construction industry, the use of three-dimensional (3D) virtual models of buildings is quickly increased. Using three-dimensional models can give people a more convenient and intuitive understanding of the model of the building, and it is necessary for the painter to manually draw the 3D model. By analyzing the common design rules of architectural drawing, this project designed and realized a building three-dimensional reconstruction system that can automatically generate a stereogram (3 ds format) from a building plan (dxf format). The system extracts the building information in the dxf plan and generates a three-dimensional model (3 ds format) after identification and analysis. Three-dimensional reconstruction of architectural drawings is an important application of computer graphics in the field of architecture. The technology is based on computer vision and pattern recognition, supported by artificial intelligence, three-dimensional reconstruction, and other aspects of computer technology and engineering domain knowledge. It specializes in processing architectural engineering drawings with rich semantic information and various description forms to automatically carry out architectural drawing layouts. The high-level information with domain meanings such as the geometry and semantics/functions of graphics of the buildings can be analyzed for forming a complete and independent research system. As a new field of computer technology, the three-dimensional reconstruction drawings are appropriate for demonstrating the characteristics of architectural constructions.

1. Introduction

As the complexity of construction project management continues to increase, more and more automated and intelligent construction schedule methods have attracted attention from the traditional management field. However, the existing mainstream methods are subject to high cost and complex use restrictions and are difficult to apply to the complex construction schedule management scene. By comparing the characteristics of various 3D reconstruction technologies, a building construction schedule collaborative management automation system (DLR-P) based on deep learning 3D reconstruction technology was built. The system uses a high-speed camera to collect real-time image information of the construction site to complete the reconstruction from two-dimensional information to three-dimensional

information and is combined with BIM dynamic model technology to achieve automatic control of the construction progress. Taking the construction site of a project in Banan District, Chongqing City as an example, the empirical study of the system was carried out, and various data during the operation of the system were verified and analyzed. The results show that the average 3D reconstruction time of the DLR-P system is 61s, which satisfies the basic schedule management requirements, can realize the automated management of the construction schedule, and effectively improves the efficiency of the construction schedule management. Compared with the existing management methods, it shows greater advantages in terms of operating costs and ease of use [1]. Construction schedule management of construction projects runs through the entire life cycle of construction [2].

Degree management efficiency is low, and it has caused cost overruns and legal disputes in many construction projects due to delays in construction schedules [3]. In order to achieve real-time, convenient, and economical construction schedule automation management, based on the existing intelligent construction field construction schedule automation management framework, a construction schedule collaborative management system framework based on deep learning 3D reconstruction technology is proposed, as shown in Figure 1, and describes the construction of the core part of the system's 3D reconstruction deep learning model and the system operation process [4, 5]. Regarding the topic of automated management of building construction schedules, scholars have conducted a lot of research in combination with various technologies. However, the existing research is difficult to apply to the complex practical requirements of building construction management [6]. The goal of the 3D reconstruction system of the building construction scene is to complete the whole parameter and whole process control of building construction, which integrates the basic functions. An important branch of artificial intelligence, known as deep learning, is so useful in smart building construction. Therefore, deep learning technology has extraordinary practical significance for the combination of artificial intelligence and the construction industry [7–9].

In recent years, the research and application of 3D reconstruction technology have been rapidly developed, but they still face many problems. To this end, this article will focus on the main progress and some representative research results of vision-based 3D reconstruction technology in recent years. Introduce, provide reference for scientific research personnel, and on this basis, through comparison and analysis, explore the difficulties and hotspots in the research of 3D reconstruction technology, as well as possible development trends [10]. On the whole, the 3D reconstruction technology mainly uses the visual sensor to obtain the real information of the outside world and then obtains the 3D information of the object through the information processing technology or the projection model. In other words, the 3D reconstruction is a way of using 2D projection to restore the 3D information computer technology [11]. With the continuous advancement of science and technology, many research directions have emerged in 3D reconstruction technology, among which construction site construction is one of the main research directions of 3D reconstruction technology [12]. Some countries in Europe and the United States have already carried out research on 3D laser scanning technology. Long ago, Stanford University has carried out research on large-scale ground-mounted laser scanning systems and obtained more accurate experimental results [13]. Architectural technology refers to the methods and means to support the construction of construction projects, and the contextualization of architectural structures is the spatial intention of buildings. Among them, the architectural technology can convey the architectural intention. Besides, the actual construction effect of the building can be matched with the preconceived scene. In summary, the 3D modeling technology has been widely used

in the field of architecture. Using 3D modeling technology to make architectural models, there is a realistic effect, in place. In addition, you can view the different effects of various parts of the building and the surrounding environment from a full range of angles. The innovation of this study lies in the adoption of 3ds Max software as a development tool to complete the three-dimensional modeling and production of buildings, the production of topographic maps, the optimization of various scene models, and the final model output. structure. Therefore, the combination of the two can make the results of construction implementation have corresponding landing and controllability [14].

2. The Important Role of Digitalization in 3D Reconstruction of Buildings

At present, there are many research studies on building 3D data acquisition and building 3D reconstruction at home and abroad, but there are many differences in their methods. According to the source of the original data, they are mainly divided into the following three categories: based on satellite remote sensing images, laser point cloud, and tilt-based photography [15]. Three-dimensional building reconstruction techniques based on oblique photography are mostly used in nonprofessional fields. The viewfinder materials for oblique photography mainly include ground cameras and unmanned aerial photography. If the area is low and the density of the building is low, the ground camera is often used; if the building is taller and the building is dense, an unmanned camera can be used to shoot from the air. Use camera materials to take multiangle photographs of buildings and record the corresponding angles of the photographs. According to the relevant image processing and data processing algorithms, the three-dimensional architectural outline data of the building surface is calculated. In order to quickly obtain an accurate three-dimensional model of the construction site, it is necessary to obtain a three-dimensional model with high accuracy, high real-time performance, and strong anti-interference. The high precision of the 3D model is mainly evaluated from two directions: high-quality plane model and high-precision camera positioning. A high-quality plane model requires a three-dimensional model of the site which has rich texture information and is noise-free. This model should simulate the current three-dimensional scene with a continuous surface rather than a discrete point cloud and is less disturbed by external noise [16]. In addition, the most basic matching algorithm in the reconstruction algorithm is the nearest point iterative algorithm, but in practical applications, this algorithm is susceptible to interference from external factors. In addition, there are matching algorithms based on feature point extraction, such as SIFT algorithm, SURF algorithm, and ORB algorithm, but there is often error accumulation in the matching process. In addition to interframe matching based on the acquired data, there are also the realization of 3D reconstruction using global pose optimization, closed-loop detection, and beam adjustment, through image or key point-based relocation algorithms, which can reduce the matching error in the reconstruction process and external noise interference, thereby improving the accuracy of 3D model reconstruction [17].

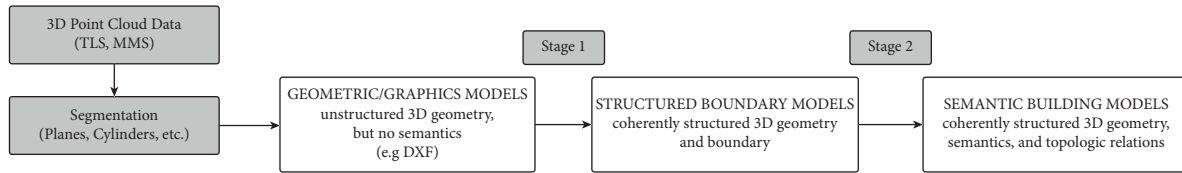


FIGURE 1: The deep learning framework with reconstruction technology.

The architectural space scene expression is the image display of the architectural space vision. Architectural construction is the practical process of implementing and constructing the scene. That is, the architectural space scene first uses the three-dimensional technology to present the scene-oriented planning and design and then use this planning and design to guide the actual construction, and the construction of the project brings the actual construction results closer to the expected effect in the design. Therefore, from architectural scene design to building construction is a connected process and an organic whole [18]. Architectural structure sceneization and building construction technology are complementary conceptual systems. The purpose of building construction technology application is to realize building structure sceneization. For example, building construction technology uses building materials to realize the relationship between architectural form design and architecture or landscape. It conveys the space effect and display effect that the building wants to express. At the same time, the sceneization of architectural structure can guide the construction of buildings through specific objects such as space display and landscape design. For example, the sceneization of existing architectural structures can inspire and even guide the construction of new architectural spaces [19]. In architectural design, the scene-based design of the building site is the embodiment of the construction results, and it is inseparable from the support of building construction technology. The different ways and themes of architectural scene expression will guide the reference of different construction techniques. For example, some scene space is relatively open, and it needs to be realized by combining corresponding construction technology and building materials in construction. For example, the content of the scene expression is comprehensive, and the building construction technology needs to be used to integrate the architectural space and the external environment of the building as a whole to form a unified scene space [20]. Deep learning is to simulate the visual mechanism of the human brain by combining low-level features to form more abstract high-level features or attribute categories to achieve complex function approximation and distributed representation of input data. The existing deep learning models mainly include convolutional neural networks, restricted Boltzmann machines, and deep belief networks. Deep learning is widely used in image recognition and behavior recognition, especially in the field of computer vision, which can realize the automatic recognition and classification of image scenes and the extraction of action gestures. Applying deep learning theories and methods can realize scene processing from the perspective of artificial intelligence, thereby effectively

improving processing efficiency, which is explained as Figure 1 shows.

People generally abstract the specific spatial relationship of a set of specific objects as a scene. The constituent factors of the scene can be summarized in three aspects: (1) nonfixed characteristic factors refer to the actors in the scene with different characteristic attributes, showing different behavioral characteristics; (2) fixed characteristic factors refer to factors that are basically fixed or change little, such as floor construction characteristic environment; (3) semi-fixed characteristics refer to even in the same space. Different locations can produce and guide the occurrence of different behaviors, thereby forming different scenarios, such as pipelines and scaffolding [21]. Analyzed from the perspective of computer vision, the image semantics of construction scenes can be divided into three levels: low-level features, middle-level semantics, and high-level vision based on different levels of understanding. (1) The underlying features are low-level visual information, which can be directly obtained from the image, which is the most direct and objective description of the visual features of the image. (2) The middle-level semantics is represented by a visual packet model or semantic topics. The middle-level semantics is derived from the low-level visual features to express the feature information that is based on the content. (3) High-level vision refers to the semantic information obtained by people's high-level abstract cognition of images, which often contains higher-level and more abstract semantics than lower levels. There is a "semantic gap" between the low-level features and the high-level vision that people understand; that is, when judging image similarity, it is based on understanding the semantics of the description object, not just based on features such as the texture, color, and shape of the bottom layer.

3. Construction Site Management System Based on 3D Reconstruction Technology

3.1. System Framework. The framework of the construction progress collaborative management system based on deep learning 3D reconstruction technology consists of the following 4 parts to manage the closed-loop [22], as Figure 2 illustrated.

- (1) Construction site information: the construction site information is used as the frame. The basic data of the framework provides a management data basis for closed-loop management. The system uses high-speed camera sensors to collect 3D information on the construction site progress and then uses the 3D

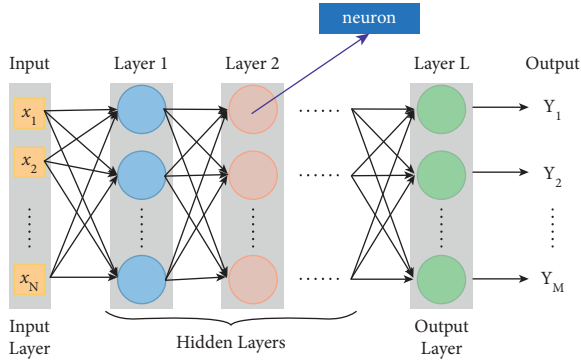


FIGURE 2: The logic diagram of the deep learning algorithm.

reconstruction deep learning model to realize the construction of a digital model of the construction site progress.

- (2) Three-dimensional reconstruction model of construction progress: based on real-time acquisition of multiview pictures, camera internal parameters, camera external parameters, and data matching information from the construction site, through features extraction, construction cost matching, depth estimation optimization, and point cloud model fusion technologies, the process obtains the actual 3D point cloud model of the construction site.
- (3) Ideal BIM construction schedule model: the model is the expected BIM construction schedule model (a 4D BIM model) jointly formulated by the design unit, the construction unit, and the owner before the construction of the building project, which includes not only the three-dimensional information during the construction process but also the progress information of the construction progress along with the time progress during the construction process. This part of the information has been formulated before the start of construction. During the construction process, this model is cross-compared with the point cloud model obtained in (2) to obtain construction progress difference information and then generate corresponding construction site resource adjustment opinions.
- (4) Construction site resource information: during the construction process, on-site management personnel, based on the construction site resource adjustment opinions generated in the process of (3), organize on-site labor, materials, machinery, and other resources to respond to achieve the purpose of on-site schedule adjustment, and after adjustment, the ideal BIM construction schedule model is dynamically adjusted to meet the overall schedule requirements.

3.2. 3D Reconstruction Deep Learning Model. The 3D reconstruction deep learning model is the core part of the system. The deep learning model used by the DLR-P system is MVSNet proposed by the Yaoyao team in 2018 [23]. This

method is a classic three-dimensional reconstruction method proposed in recent years. While achieving good reconstruction results, it has also been used as an extension of the basic model to develop a series of deep learning models. Model principle: the depth information of different spatial positions is merged to construct the surface three-dimensional model information of the object. Model structure: the MVSNet model structure according to its functions mainly includes three parts: feature extraction, construction matching cost, depth estimation, and optimization.

- (1) Feature extraction refers to image features extracted by a neural network. After the viewing angle is selected, the paired images, namely, the reference image and the candidate set, are input to the network model, an 8-layer two-dimensional convolutional neural network is used to extract the depth features of the stereopair, and the 32-channel feature map is output. In order to prevent the loss of semantic information after the input image is downsampled, the semantic information between the neighboring pixels of the pixel has been encoded into the 32-channel feature, and the network of each image extraction process is weight-sharing.
- (2) Constructing the matching cost: the model uses the plane scanning algorithm to construct the matching cost of the reference image. After the feature extraction process, each image can obtain a corresponding feature map. According to the prior experience depth range, the reference image is scanned in the direction of its main optical axis, and the reference image is scanned from the minimum depth to the maximum according to a certain depth interval. Depth mapping: you can get a camera cone at different depth intervals, as shown in Figure 2. The feature maps in the candidate set are mapped to the camera cone. Through projection transformation, several images can form a corresponding number of feature bodies. This feature body is a representation of the matching cost. Finally, the cost accumulation of MVSNet is realized by constructing a three-dimensional structure composed of a cost map whose length and width are the same as the length and width of the reference image in the depth direction [24, 25].

As shown in the following formula:

$$E(u, v) = \sum_{x, y}^n w(x, y) [I(x + u, y + v) - i(x, y)]^2, \quad (1)$$

where u and v are the offset

- (3) Depth estimation and optimization: MVSNet's depth estimation is obtained by direct learning through deep neural networks. The network training method is to input the cost body and the corresponding true value of the depth map and use the SoftMax function to return the probability value of each pixel at the

depth θ to complete the learning process from the cost to the depth value. The depth map and RGB image generated by the final model can be fused into a point cloud model. The algorithm hierarchy diagram is shown in Figure 3. When building the framework based on building information, this paper adopts the collaborative filtering algorithm to accelerate the data calculation speed, reduce the real-time response time, and enhance scalability. Thus, the accuracy, grammar, and real time of reconstruction levels are improved.

3.3. DLR-P System Operation Process. The main operating steps of the construction progress collaborative management system based on deep learning 3D reconstruction technology are as follows:

- (1) Ideal BIM model construction: before running the DLR-P system, the construction should be constructed based on the project task objectives, engineering characteristics, and project environment. The collaborative design BIM model of the project greatly improves the efficiency of collaborative design of multiple specialties in a project at the same time. Besides, two-dimensional drawings are generated by direct projection and cutting of three-dimensional models, which greatly reduces the drawing workload of designers. The model should include the three-dimensional information of the project, expected progress information, expected cost information, labor demand information, material entry and exit information, and mechanical equipment demand information.
- (2) Cooperative system construction: the system construction mainly includes two parts: an information collection module and a background processing module. The information collection module refers to the need to place the camera sensor in the required position of the construction site according to the requirements of different projects, and several sensors form an array to collect real-time appearance data of various construction site progress control targets. The background processing module includes a data processing part composed of a high-performance computer group and a progress management graphical part composed of a high-definition display. The communication between the above two modules is realized by a wireless local area network connection.

$$\text{loss} = \frac{1}{2m} \sum (\text{prediction}^{(i)} - t^{(i)})^2. \quad (2)$$

This formula represents the definition of the loss function.

- (3) System operation: make sure that after each part of the system is built, link each part of the system in the same local area network environment, and set the image sensor capture angle and capture cycle. As the

construction progress develops, ensure that the system runs in real time and that the sensor group can capture target images at different appropriate angles. The image data collected at the construction site is transferred to the background via the wireless network. Besides, the data set with a noise feature of 3D information is given in Figure 4. First, Colmap software is used to perform sparse reconstruction to calculate the camera pose matching information and other data, and then, the MVSNet 3D reconstruction deep learning model is used to generate the point cloud model of the corresponding scene. The system backstage imports the aforementioned point cloud model into the Revit software and the ideal BIM model for size comparison and compares and calculates the point cloud model and the ideal BIM model schedule according to the construction progress. Finally, the system outputs the current progress status and the corresponding construction site in a graphical display. Regulatory opinions can realize the fine management throughout the construction industry safety management [26–28].

$$\text{loss} = \frac{1}{2m} \sum (\text{prediction}^{(i)} - t^{(i)})^2, \quad (3)$$

$$L(p, u, t_u, v) = L_{cls}(p, u) + \lambda[u \geq 1]L_{ioc}(t_u, v),$$

where N is the number of index categories, and t is the displacement in the logarithmic space.

According to the control opinions output by the DLR-P system, the personnel arrangement, material entry and exit, and the use of mechanical equipment for each relevant process on the construction site are uniformly deployed to optimize the construction schedule management.

- (4) Model adjustment: input the active adjustment information such as personnel arrangement, material entry and exit, use of mechanical equipment on the construction site into the ideal BIM model, and adjust and optimize subsequent project construction work according to project requirements such as construction period and cost to form a system. The internal information of timely feedback is designed as a closed loop.

4. Image Semantic Extraction

Use CNN's deep learning structure to extract object semantics, including four processes of information input, preprocessing, feature extraction and selection, and classification decision learning. CNN can directly process two-dimensional images. Image feature extraction and dimensionality reduction are carried out step by step (Figure 5). The convolutional layer extracts image features through convolution kernels, the sampling layer reduces the dimensionality of image features, and the fully connected layer and classification layer are used for classification.

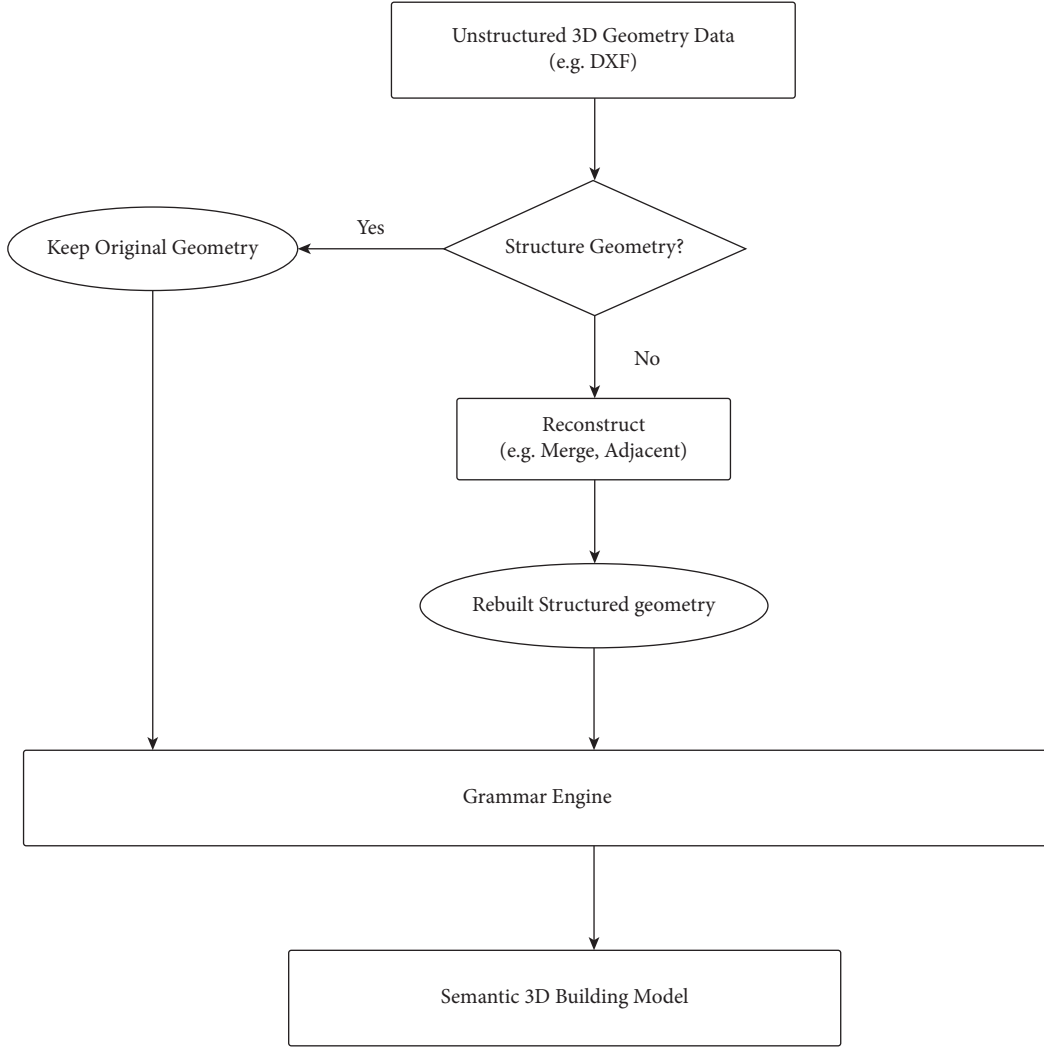


FIGURE 3: The algorithm hierarchy diagram.

The histogram of gray values at different stages is given in Figure 6. In Figure 6, the target gray value on the y-coordinate represents the gray value of the reconstruction of the 3D building model expected by the model. Furthermore, the abscissa represents the different stages of the model refactoring process. The main implementation process of object semantic extraction is as follows:

- (1) A 9×9 convolution kernel is used to convolve the input image, and the activation value of each neuron is mapped between -1 and 1 through the activation function \tanh , and a total of 32 feature maps are output. Each output map may be the value of multiple input maps combined and convolved, specifically expressed as [29]

$$X_j^l = f\left(\sum_{n \in M_j} X_j^{l-1} \times k_{nj}^l + b_j^l\right), \quad (4)$$

where f is the activation function \tanh ; X is the output of the j -th image in the first layer; M_j is the set

of input maps; k is one of the j -th images in the first layer and the n -th image in the 1-1 layer weight of the interval; b is the bias term of the j -th image in the first layer.

- (2) Use the obtained 32 feature maps as the input of the second sampling layer, and perform dimensionality reduction processing. For this subsampling layer, there are 32 input maps and 32 output maps, and both dimensions are reduced to half of the original, which is specifically expressed as

$$X_j^l = f(\text{down}(X_j^{l-1}) + b_j^l), \quad (5)$$

where down represents the downsampling function.

- (3) The output image of the second sampling layer is used as the input of the third convolution. Convolve the input image with a 9×9 convolution kernel, and use equation (1) to obtain 96 feature images. The coordinate relationship established by two related coordinate systems is as follows:

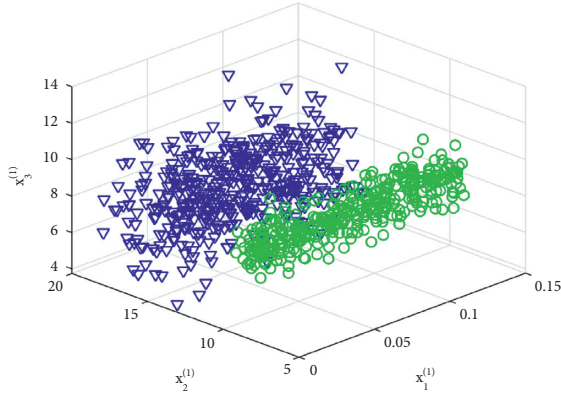


FIGURE 4: The data set with a noise feature of 3D information.

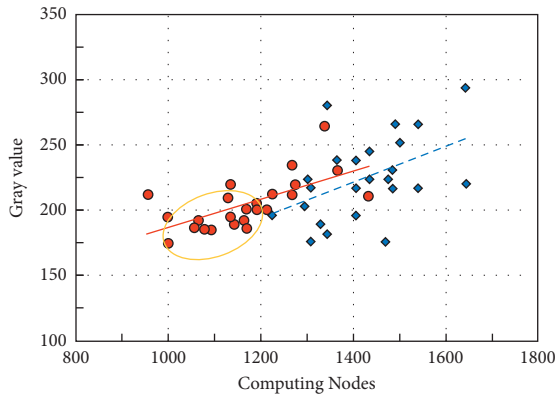


FIGURE 5: Comparison of experimental results of deep learning model.

$$\begin{aligned} u &= \frac{x}{d_x} + u_0, \\ v &= \frac{y}{d_y} + v_0, \end{aligned} \quad (6)$$

where u and v are the axis corresponding to the two coordinate systems.

- (4) The fourth subsampling layer again uses equation (2) to reduce the dimensions of the obtained convolutional features and obtain 96 10×10 feature maps, which are fully connected through the hidden layer and mapped to $-1 \sim 1$ through the tank function between them. The logistic classifier is used for identification and classification, and the results are stored in the form of a variable to provide information support for the subsequent identification of the positional relationship between objects [30].

$$\begin{aligned} x &= \frac{fX_c}{Z}, \\ y &= \frac{fY_c}{Z}, \end{aligned} \quad (7)$$

where $X Y Z$ is the coordinate of point P .

For the semantics of spatial relations, on the basis of object recognition, in order to locate each object area, the minimum edge rectangle method is used to approximate the area. The processed result is a series of rectangles, and each rectangle corresponds to a closed curve of the object contour. The error between calculation and the actual measurement is illustrated in Figure 7. The azimuth relationship between objects can be determined by establishing a directional relationship matrix.

5. Scene Data Processing

Scenario data can be expressed as the internal information of data mining from multiple scenarios and multiple dimensions, and the curve of calculation times and sample error is shown in Figure 8. Combined with existing research results, seven dimensions of unsafe behavior pan-scene data are determined from the perspective of manual collection, namely, time, location area, individual behavior, unsafe actions, nature of the behavior, traces of behavior, and level of risk. Unsafe state scenes can be described in two dimensions: scene objects and unsafe states. From the perspective of automated data collection, image semantic information is processed by multidimensional coding and relying on industry databases to form multidimensional pan-scene data. Obtained directly from digital devices, behavior traces and risk levels can be obtained from industry databases based on behavior semantics.

Formula (8) is used to quantify the change of the grayscale of the physical quantity of the local image.

$$\begin{aligned} E(x, y) &= \sum w_{uv} (I_{x+u, y+v} - I_{u, v})^2 \\ &= \sum w_{uv} \left[x \frac{\partial I}{\partial x} + \frac{\partial I}{\partial y} + o(x^2 + y^2) \right]^2. \end{aligned} \quad (8)$$

Take the scenario of unsafe behavior-workers sleeping on a construction platform as an example, and specify the corresponding relationship: based on deep learning, the object semantics is that the worker wears a yellow helmet, and the corresponding behavior individual is the grassroots worker; spatial relationship semantics, and scene. The semantic common description is that the worker is on the construction bench, and the corresponding location area is the construction bench; the behavior semantics is that the employee sleeps on the construction bench, and the corresponding unsafe action is sleeping on the construction bench; based on the industry database, it is unsafe. The nature of the action behavior is an illegal action, the behavior trace is no trace, and the risk level is average. The average number of iterations and the median number of iterations are shown in Figure 9.

Applying deep learning methods to process construction site inspection pictures can identify unsafe human behaviors and unsafe conditions of objects. At the same time, it can process construction safety pan-scene data based on image semantics in terms of the application of construction safety pan-scenario data. As Figure 10 shows, the value of mean

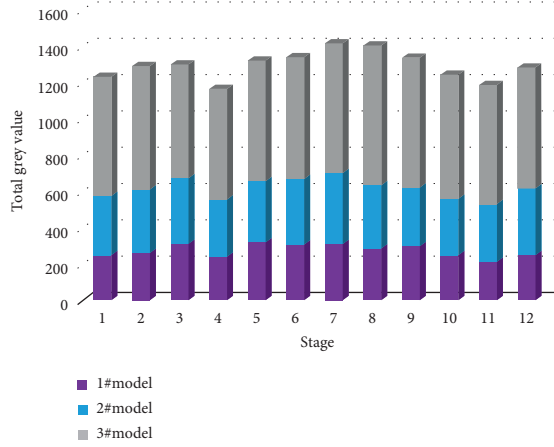


FIGURE 6: Histogram of gray values at different stages.

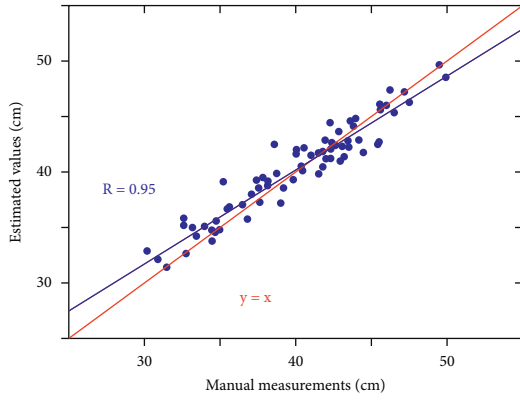


FIGURE 7: Error between calculation and actual measurement.

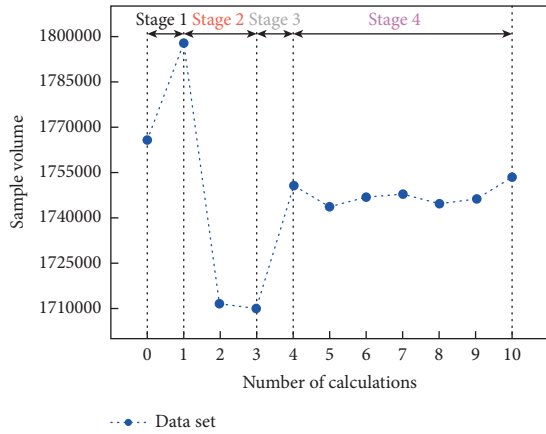


FIGURE 8: The curve of calculation times and sample error.

square error predicted by the C-means clustering algorithm of the physical quantity in local images fluctuates as the number of iterations.

- (1) Statistical analysis of different dimensions can reflect the temporal and spatial distribution of unsafe behaviors, individual behavior classification, unsafe actions summary, and other internal knowledge, so

as to provide enterprises with an accurate understanding of the characteristics of unsafe behaviors.

$$F = U_2^T F U_1 = U_2^T [e]_v U_1. \quad (9)$$

Among them, the basic matrix F can be decomposed to obtain the transformation matrices U_1 and U_2 ,

- (2) Explore the in-depth and multidimensional association rules of unsafe behaviors, and analyze when and what unsafe actions are prone to occur, where and what unsafe actions are prone to occur, and what kind of people are prone to what unsafe actions, so as to achieve personalized interventions on unsafe behaviors. Therefore, the 3D reconstructing technology can direct the process of architectural constructions in detail.

6. Advantages of Using 3D Reconstruction System in Building Construction Scenes

In the case study of the DLR-P system, the collaborative management method of construction progress based on deep learning 3D reconstruction technology showed the following three significant advantages:

- (1) Automatic and efficient: compared with the traditional construction schedule management method that relies on manual patrol on the construction site, the proposed management method greatly improves the management efficiency and realizes the unmanned operation of the whole process, thereby reducing the labor intensity of the on-site management staff and achieving more objectiveness. The on-site progress monitoring refers to the management of the progress of each stage and the deadline of the final completion of the project in the process of project implementation. It is to draw up a reasonable and economical schedule within the specified time. The on-site progress monitoring eliminates various influencing factors in the manual management process.

$$\sum_{i=1}^n \left[\frac{u_3^T (p_i - p_c)}{u_3^T p_c} \right]^2 = \frac{u_3^T P P u_3}{u_3^T p_c p_c^T u_3}, \quad (10)$$

where λ represents the scale factor, w_i is called the weight, and u_3 is the last line of u_p .

- (2) Cost economy: compared with the existing automatic management methods of “UAV + BIM” and “LiDAR + BIM,” the proposed management method greatly reduces the cost of system deployment and has higher promotion and practical value. In addition, while controlling the low cost of the system itself, this method can also improve the economic benefits of construction projects from the perspectives of saving labor costs, reducing rework waste caused by construction schedule factors, and optimizing the configuration of personnel, materials, and machinery.

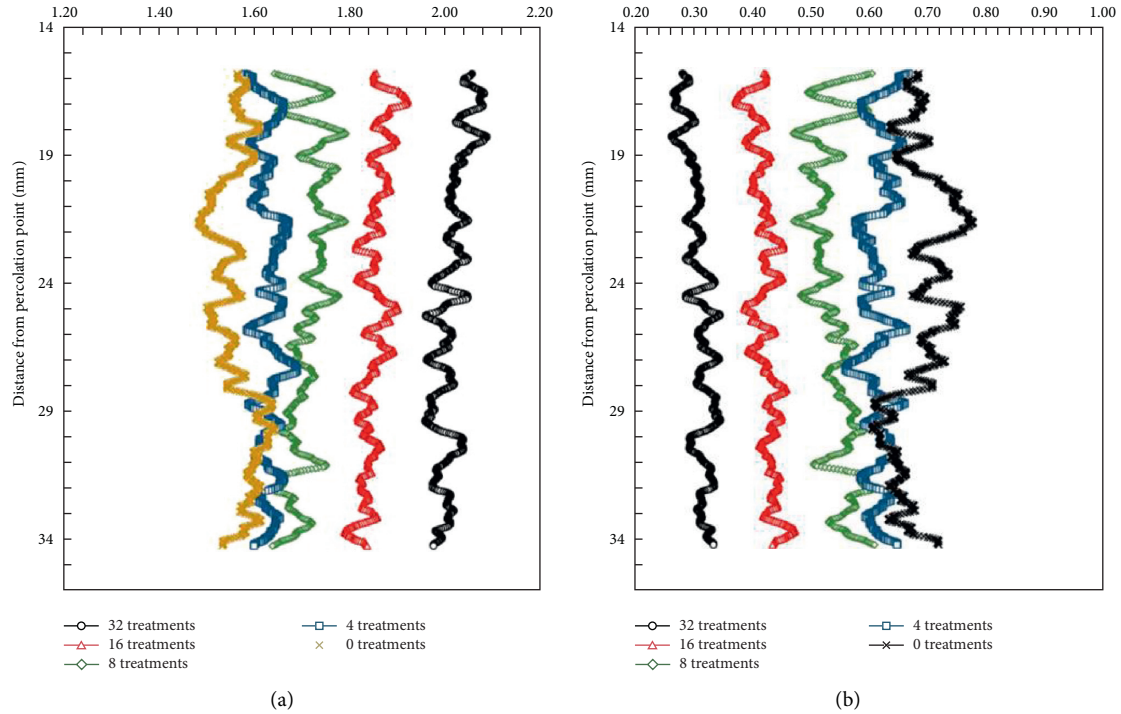


FIGURE 9: The average number of iterations and the median number of iterations. (a) The average number of iterations. (b) The median number of iterations.

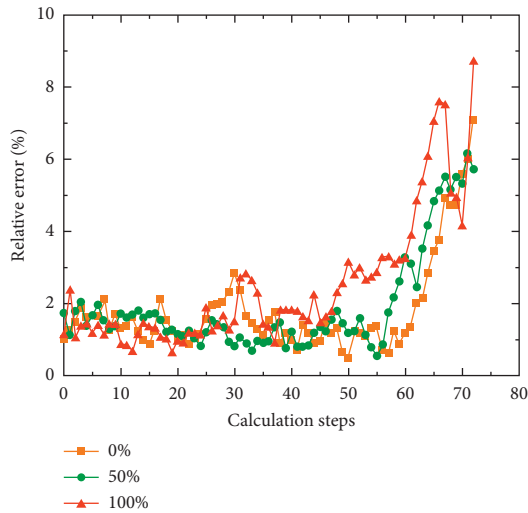


FIGURE 10: The value of mean square error predicted by the C-means clustering algorithm.

$$g(z) = \frac{t^T D^{-T} B D^{-1} t}{t^T t}. \quad (11)$$

Among them, A is symmetric and positive definite, so A can be decomposed into $A = D^T D$. Let $t = Dz$.

- (3) Convenient application: the progress management of a construction project is a work that runs through the entire stage of the project construction. It has two important characteristics: long-term and dynamic. Therefore, the automated and intelligent way of

realizing schedule management must be applicable to all kinds of changes in construction. Work in the scene: the automated schedule control system using high-speed cameras as the data collection method is more convenient to use than the method based on drones or LiDAR equipment in the process of project schedule management and control.

It is necessary to train drone operators and set up complex circulation routes in order to avoid collisions, and there is no need to consider the limitations of scanning instruments in special locations on the spot, which greatly liberates the information collection capabilities of the system and minimizes the schedule management work. The impact on various construction procedures is shown.

7. Conclusion

The 3D reconstruction technology based on deep learning shows good economic efficiency in the management of construction progress. Compared with the UAV method or lidar method in the current mainstream research, since only a high-speed camera is required as a sensor to collect data, it has significant advantages in terms of equipment cost, operating labor cost, and on-site coordination cost. The automation and intelligent management of building construction progress runs through the entire construction cycle, and the economic characteristics of this method can better meet the cost-sensitive characteristics of construction management units.

On the basis of extracting image object semantics, spatial relationship semantics, scene semantics, and behavior

semantics, the corresponding relationship between semantic information and pan-scene data is constructed, and then through the industry unsafe behavior and unsafe physical state database, the automation of pan-scene data can be realized to deal with.

As the 3D reconstruction technology of building construction scenes is not very mature at present, many technical research studies are currently in the research stage. Due to various reasons, there has not been a widely used 3D reconstruction system of architectural drawings on the market so far. The 3D reconstruction system of architectural drawings realized by this system has some characteristics and some limitations. It can be further improved and perfected. It is hoped that a powerful, easy-to-operate, and widely used 3D reconstruction system of architectural drawings can be realized. Serving as a new technology, the three-dimensional drawings with applying reconstructing technology with the computer are appropriate for demonstrating the detailed geometric features of architectural constructions.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] S. Yang, P. Luo, C. Loy, and X. Tang, "From facial parts responses to face detection: a deep learning approach," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3676–3684, Santiago, Chile, 7–13 Dec. 2015.
- [2] J. Chen, J. Wu, and Y. Qu, "Monitoring construction progress based on 4D BIM technology," *IOP Conference Series: Earth and Environmental Science*, vol. 455, no. 1, Article ID 012034, 2020.
- [3] T. H. Kim, W. Woo, and K. Chuang, "3D scanning data coordination and as-built-construction process optimization utilization of point cloud data for structural analysis," *Architectural Research*, vol. 21, pp. 111–116, 2019.
- [4] L. Yang, J. C. P. Cheng, and Q. Wang, "Semi-automated generation of parametric BIM for steel structures based on terrestrial laser scanning data," *Automation in Construction*, vol. 112, Article ID 103037, 2020.
- [5] K. Zhou, R. Lindenbergh, and B. Gorte, "LiDAR-guided dense matching for detecting changes and updating of buildings in Airborne LiDAR data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 162, pp. 200–213, 2020.
- [6] M. Ali, L. Corti, F. Delleani, and D. Zanoni, "3D reconstruction of fabric and metamorphic domains in a slice of continental crust involved in the Alpine subduction system: the example of Mt. Mucrone (Sesia-Lanzo Zone, Western Alps)," *International Journal of Earth Sciences*, vol. 109, no. 4, pp. 1337–1354, 2020.
- [7] J. Chi, J. Liu, F. Wang, Y. K. Chi, and Z. G. Houg, "3D gaze estimation method using a multi-camera-multi-light-source system," *IEEE Transactions on Instrumentation and Measurement*, vol. 99, p. 1, 2020.
- [8] C. Sung and P. Y. Kim, "3D terrain reconstruction of construction sites using a stereo camera," *Automation in Construction*, vol. 64, no. APR, pp. 65–77, 2016.
- [9] L. Hoegner, S. Tuttas, and U. Stilla, "3D building reconstruction and construction site monitoring from RGB and TIR image sets," in *Proceedings of the 2016 12th IEEE International Symposium on Electronics and Telecommunications (ISETC)*, 27–28 Oct. 2016.
- [10] R. Tong, C. Chen, and P. Cui, "Deep learning method for processing pan-scene data on construction safety," *China Safety Science Journal*, vol. 27, 2017.
- [11] F. Yu, Y. Zhang, and S. Song, "LSUN: construction of a large-scale image dataset using deep learning with humans in the loop," *Computer Science*, <https://arxiv.org/abs/1506.03365>, 2015.
- [12] Y. Zhao, Q. Chen, and W. Cao, "Deep learning for risk detection and trajectory tracking at construction sites," *IEEE Access*, vol. 7, p. 1, 2019.
- [13] M. Arashpour, T. Ngo, and H. Li, "Scene understanding in construction and buildings using image processing methods: a comprehensive review and a case study," *Journal of Building Engineering*, vol. 33, Article ID 101672, 2020.
- [14] C. Wang, Q. Yu, and K. H. Law, "Machine learning-based regional scale intelligent modeling of building information for natural hazard risk management," *Automation in Construction*, vol. 122, 2021.
- [15] S. Y. Wen and A. Y. Chen, "Tracking multiple construction workers through deep learning and the gradient based method with re-matching based on multi-object tracking accuracy," *Automation in Construction*, vol. 119, Article ID 103308, 2020.
- [16] W. U. Zhen, H. Y. Liu, and L. I. Xue, "Research on precise 3D model reconstruction of tall and complicated building," *Journal of Henan University of Urban Construction*, vol. 25, no. 2, pp. 63–67, 2016.
- [17] Z. Chen, Y. Zhang, and J. Li, "Diagnosing tunnel collapse sections based on TBM tunneling big data and deep learning: a case study on the Yinsong Project, China," *Tunnelling and Underground Space Technology*, vol. 108, Article ID 103700, 2020.
- [18] Y. Chang, R. J. Ries, and Y. Wang, "Life-cycle energy of residential buildings in China," *Energy Policy*, vol. 62, pp. 656–664, 2013.
- [19] T. Ninomiya, M. Sugimura, Y. Yoshino, and T. Komuro, "Design and construction of the super high-rise building with deep underground structure," *Concrete Journal*, vol. 55, 2017.
- [20] H. S. Cha and J. Kim, "A study on 3D/BIM-based on-site performance measurement system for building construction," *Journal of Asian Architecture and Building Engineering*, no. 1, pp. 1–12, 2020.
- [21] X. Zhang, W. Liu, Y. Li, S. Yang, and X. Liu, "Structural selection and design analysis of a super high-rise office building," *IOP Conference Series: Earth and Environmental Science*, vol. 510, no. 5, Article ID 052004, 2020.
- [22] J. Chen, Z. Kira, and Y. K. Cho, "Deep learning approach to point cloud scene understanding for automated scan to 3D reconstruction," *Journal of Computing in Civil Engineering*, vol. 33, no. 4, pp. 04019027.1–04019027.10, 2019.
- [23] T. Czerniawski and F. Leite, "Automated segmentation of RGB-D images into a comprehensive set of building components using deep learning," *Advanced Engineering Informatics*, vol. 45, Article ID 101131, 2020.
- [24] M. Pashaei, M. J. Starek, H. Kamangir et al., "Deep learning-based single image super-resolution: an investigation for

- dense scene reconstruction with UAS photogrammetry,” *Remote Sensing*, vol. 12, no. 11, p. 1757, 2020.
- [25] L. Wang, Z. Huang, Y. Gong, and C. Pan, “Ensemble based deep networks for image super-resolution,” *Pattern Recognition*, vol. 68, pp. 191–198, 2017.
- [26] J. Q. Zhang, C. L. Wu, and J. Q. Fan, “The research of landslide monitoring data integration framework based on three-dimensional webgis,” *Applied Mechanics & Materials*, vol. 694, pp. 436–441, 2014.
- [27] C. A. Tryon, J. T. Faith, D. J. Peppe, E. J. Beverly, N. Blegen, and S. A. Blumenthal, “The pleistocene prehistory of the lake victoria basin,” *Quaternary International*, vol. 404, pp. 100–114, 2016.
- [28] N. Chen, X. U. Zhao, and L. I. Ming, “A BIM-based study on environmental impacts and decision making analysis of building construction,” *Journal of Engineering Management*, vol. 30, no. 2, pp. 97–102, 2016.
- [29] Y. Gan, H. Wang, and Y. Peng, “Building and application of construction information model of converter station’s site formation based on BIM and 3D GIS technology,” *Geotechnical Investigation & Surveying*, vol. 41, 2019.
- [30] B. Taylor, A. K. Jones, and J. Yang, “Measuring computed tomography scanner variability of radiomics features,” *Investigative Radiology*, vol. 50, no. 11, pp. 757–765, 2015.

Research Article

Dynamic Analysis of Deep Water Highway Tunnel under Ocean Current

Li Fang , Hong Li , and Bin Li

School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan 430063, China

Correspondence should be addressed to Hong Li; lhwt@whut.edu.cn

Received 28 January 2022; Revised 21 February 2022; Accepted 8 March 2022; Published 31 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Li 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.

Comprehensively comparing the merits and demerits of the existing means of transportation across the water, a new underwater transportation structure for crossing the wide water area, named as “deep water highway tunnel” (hereinafter called “DWHT”), is proposed. The characteristics of flow field around the typical section of DWHT at different flow velocities are investigated, which can provide reference for the values of hydrodynamic coefficient at high Reynolds number. The vibration modes and natural by the sound-solid coupling method. In addition, considering the factors of fluid-structure coupling, the dynamic response of displacement and internal force is analyzed based on CFD for the weak parts of the structure. The results show that the deepening of water and the increase of flow will significantly increase the flow field pressure and structure stress, and when the span (or width-span ratio) of the tunnel body extends beyond a certain range, the dynamic characteristics and dynamic response rules of the structure will change.

1. Introduction

The core problem of underwater tunnel construction is to ensure the stability and safety of the tunnel structure system in harsh marine environment, accurately predict the hydrodynamic response of the structure system, and obtain the technical parameters such as motion, deformation, and stress, which are the primary scientific and technological problems faced by the design, construction, and safe service. At present, scholars have made some achievements in the research about the underwater structure under wave and flow load.

Svein et al. [1] proposed an alternative method for random dynamic response analysis of SFT under wave loads. Hiroshi et al. [2] conducted two-dimensional model tests for the wave environment of Hokkaido, and the dynamic characteristics of SFT are studied by boundary element numerical simulation. Di Pilato et al. [3] chose Messina channel SFT as the research object, considering the coupling vibration effect of pipe body and anchor cable, and studied its nonlinear time-history response under the action of steady flow and wind-induced waves, as well as the local

dynamic behavior of anchor cable under the action of earthquake. Tariverdilo et al. [4] established two-dimensional and three-dimensional models to simulate the dynamic response of suspended tunnel under moving load, considering the inertia effect of fluid. The results showed that the response values of the two models and their differences gradually decreased with the increase of tension leg stiffness. Seo et al. [5] carried out physical model tests in two-dimensional wave tanks for accurately estimating the hydrodynamic force and proposed a pendulum model used to describe the motion of SFT under single vertical mooring.

Mai et al. [6] took the lead in exploring the static and dynamic response of SFT under the combined action of wave and flow loads and focused on the analysis of the dynamic response of vortex-induced vibration. Wang et al. [7] and Luo et al. [8] respectively used RNG model and LES modal to analyze the pressure distribution characteristics around different sections and the variation rules of hydrodynamic parameters. Li et al. [9] studied the effects of random irregular waves on the underwater tunnel in a two-dimensional wave flume experiment. Fan and Yuan [10] explored the influence of different design parameters on the dynamic

response of the underwater tunnel and also studied on the vortex-induced resonance characteristics based on the differential equation of vibration including nonlinear factors; the results show that choosing a feasible span length was the most effective measure to control the vortex-induced resonance response. Yang et al. [11] researched on vehicle partial load effect, with a tunnel tube with a total span of 1 km as the research object, based on Hamilton principle.

Compared with the existing transportation means across water, this paper proposes a new underwater transportation structure, named as “deep water highway tunnel.” As a matter of experience about underwater tunnels, several researches on the dynamic analysis of the DWHT in the ocean current are conducted, including a series of numerical simulations based on fluid-structure coupling methods to investigate the characteristics of flow field around typical sections, as well to explore the mechanism and relevant influence factor of natural vibration characteristics and dynamic response of deep water highway tunnel.

2. A Concept: Deep Water Highway Tunnel

Researches on new means of transportation across water areas have been gradually emerging. In terms of the alternative modes to cross broad water areas, the main transportation modes include ferry, bridge, and underwater tunnel. In recent years, underwater tunnel has attracted attention for its unique advantages.

Ferry, as the most traditional modes of cross-sea transportation mode, has numerous advantages of mature technology, low construction cost, large transportation volume, and so on. However, it is easy to be affected by climate, and the fuel consumption is large, which makes ferry not suitable for the fast pace of modern urban transportation. As for the cross-sea bridge, it is convenient for vehicles to pass, with low operating costs and strong capacity. Nevertheless, there are also various problems such as geographical limitations, vulnerability to adverse weather, high cost, and difficulties of underwater construction.

Underwater tunnels are divided into undersea tunnels, immersed tunnels, and submerged floating tunnels according to the depth of the pipeline axis. The surrounding rock environment of undersea tunnel is considerably complex, leading to a challenge to carry out geological investigations. Furthermore, the long construction period, the staggering cost of the undersea tunnels, and frequent accidents such as landslides and water gushing are actually hard nuts to crack. Immersed tunnel has higher requirements on the flatness of submarine topography and limits the maximum buried depth. The construction technology of foundation groove excavation and foundation treatment is complicated, which will also cause certain damage to the marine ecological environment. Submerged floating tunnel (i.e., SFT) [12, 13] is a new type of underwater technology, although there are no completed projects of SFT at present. Especially when anchor cable is used as the support system, the cost is high, underwater installation is complex, and vortex-induced resonance is easily induced which severely limits the horizontal spacing of anchor system, and

numerous other technical difficulties are still urgently needed to be settled.

The conception of deep water highway tunnel is based on the structural model of the cross-sea bridge and the submerged floating tunnel. As shown in Figure 1, its basic structure is composed of a completely closed tunnel pipe for carrying traffic loads and a fixed support system embedded in the seabed in the form of piers and columns.

The whole body is in a certain depth under water. While the superstructure is built in water, the huge buoyancy of water can balance part of the dead weight of the structure. As meantime, the form of pier column as the supporting system in the substructure shows its own superiority. The pier column supporting system not only has stronger supporting stiffness that can bear more dynamic load but also can provide an omnidirectional restrain, both horizontal and vertical displacement of the upper tunnel pipe body are effectively constrained, which can improve the structural stability of the tunnel body.

In addition, from the perspective of engineering construction, the latest research on fixed supported underwater structures is extensive and abundant, and the relevant technologies and experience are relatively sufficient, which indicates a distinct feasibility of DWHT.

As a new traffic infrastructure, DWHT has its particularity and limitation, through a comprehensive comparison of different means of transportation across water, DWHT stands out in geographical condition, economy, environmental protection, application, and other aspects, and the specific advantages of DWHT are shown in Figure 2, whereas limited to some environmental factors, as well as the construction technology, and the pier, as a supporting system, will cause the increase of structure weight, therefore the deep water highway tunnel is more suitable for a wide marine area, where the seabed soil is relatively more bearable.

3. CFD Theory and Method

3.1. Computational Fluid Dynamics Governing Equations. The basic governing equation including the N-S equation of incompressible viscous fluid and the continuity equation are as follows:

$$\begin{cases} \rho \frac{du}{dt} = -\frac{\partial p}{\partial x} + \mu \nabla^2 u + f_x, \\ \rho \frac{dv}{dt} = -\frac{\partial p}{\partial y} + \mu \nabla^2 v + f_y, \\ \rho \frac{dw}{dt} = -\frac{\partial p}{\partial z} + \mu \nabla^2 w + f_z, \end{cases} \quad (1)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0,$$

where u, v, w are respectively the velocity components in the $X, Y,$ and Z directions; f_x, f_y, f_z are respectively the external force components of fluid per unit volume in the $X, Y,$ and Z directions.

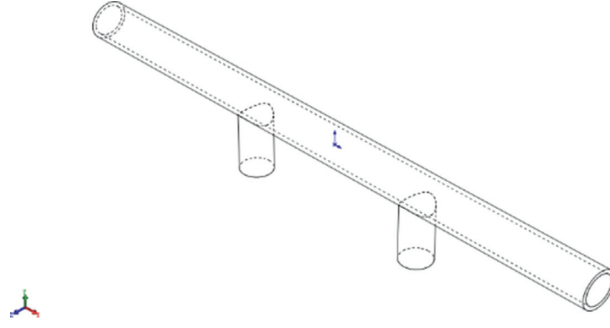


FIGURE 1: Basic structure diagram of DWHT.

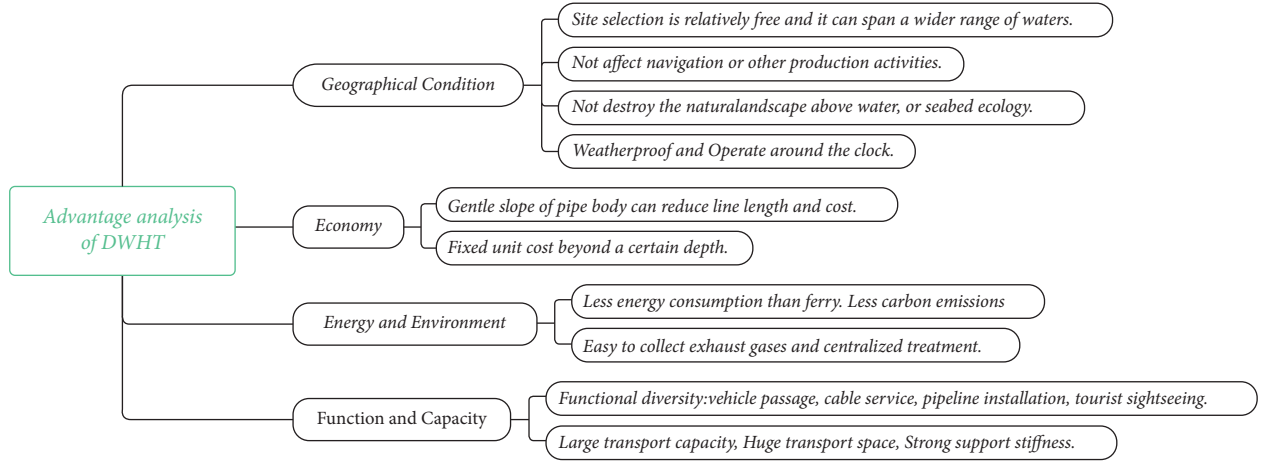


FIGURE 2: Advantages of deep water highway tunnel.

Among various turbulence models, SST $k - \omega$ model can obtain more accurate results when dealing with boundary layer simulation and complex flow separation prediction at high Reynolds number, and its governing equation is shown as follows:

$$\begin{aligned} \frac{\partial}{\partial t} (\rho k) + \frac{\partial}{\partial x_i} (\rho k u_i) &= \frac{\partial}{\partial x_j} \left(\Gamma_k \frac{\partial k}{\partial x_j} \right) + G_k - Y_k + S_k, \\ \frac{\partial}{\partial t} (\rho \omega) + \frac{\partial}{\partial x_i} (\rho \omega u_i) &= \frac{\partial}{\partial x_j} \left(\Gamma_\omega \frac{\partial \omega}{\partial x_j} \right) + G_\omega - Y_\omega + D_\omega + S_\omega, \end{aligned} \quad (2)$$

where G_k is the kinetic energy of turbulence, G_ω is the generation equation of ω , Γ_k, Γ_ω are respectively the effective diffusion terms of k and ω , Y_k, Y_ω are respectively the divergent item of k and ω , and D_ω is the quadrature divergence term of ω .

3.2. Fluid-Structure Coupling Theory. The interaction between seawater and underwater structures is dynamic. On the one hand, the fluid force causes the deformation and movement of the underwater tunnel; on the other hand, the movement of the tunnel structure changes the relative position and velocity between tunnel structure and the fluid, thus changing the fluid force. Therefore, it is necessary to consider the fluid-structure interaction effect.

The acoustic-solid coupling method is used to simulate the fluid-structure interaction in wet modal analysis of the structure. The fluid is regarded as a compressible non-viscous acoustic medium. The fluid action is presented in the form of sound pressure dynamic load on the structure wall, and the sound field and velocity potential have to satisfy the Helmholtz wave equation, which is shown as follows:

$$\frac{1}{C^2} \frac{\partial^2 P}{\partial t^2} - \nabla^2 P = 0. \quad (3)$$

Acoustic finite element discretization equation is

$$[M_a] \{\ddot{p}\} + [C_a] \{\dot{p}\} + [K_a] \{p\} + \rho [R] \{\ddot{u}\} = 0, \quad (4)$$

where $[M_a]$ is the fluid mass matrix, $[C_a]$ is the fluid damping matrix, $[K_a]$ is the fluid stiffness matrix, $[R]$ is the structural-sound field coupling matrix, $\{p\}$ is the sound pressure vector of fluid element node, ρ is the fluid density, and $\{u\}$ is the node displacement vector.

Dynamic equation of the structure is

$$[M_s] \{\ddot{u}\} + [C_s] \{\dot{u}\} + [K_s] \{u\} - [R]^T \{p\} = \{F_s\}, \quad (5)$$

where $[M_s]$ is the structure mass matrix, $[C_s]$ is the structure damping matrix, $[K_s]$ is the structure stiffness matrix, $[R]^T$ is the transpose of the structure-sound field coupling matrix, and $\{F_s\}$ is the external load vector received by the structure.

Simultaneous equations are obtained by considering the formulas above, and the acoustic-solid coupling finite element equation is shown as follows:

$$[M]\{\ddot{x}\} + [C]\{\dot{x}\} + [K]\{x\} = \{F\}, \quad (6)$$

$$\text{where } [M] = \begin{bmatrix} [M_s] & 0 \\ \rho[R] & [M_a] \end{bmatrix}, \quad \{x\} = \begin{Bmatrix} u \\ p \end{Bmatrix}, \quad \{F\} = \begin{Bmatrix} F_s \\ 0 \end{Bmatrix}, \\ [C] = \begin{bmatrix} [C_s] & 0 \\ 0 & [C_a] \end{bmatrix}, \quad \text{and } [K] = \begin{bmatrix} [K_s] & -[R]^T \\ 0 & [K_a] \end{bmatrix}.$$

As for the study of dynamic response analysis, the sequential coupling method is used to solve the problem. The most basic principle of fluid-structure coupling is that stress, displacement, and other variables should be equal or conserved on the fluid-structure coupling surface. According to the set order, the results of fluid domain and structure domain are exchanged by the fluid-solid interface, and the fluid control equation and structure control equation are solved in different solvers in turn.

4. Numerical Simulation

Dynamic research of DWHT under ocean current is carried out from three aspects: firstly, the analysis of flow field around circular section under different flow velocity. Secondly, the wet modal analysis of typical structures based on the acoustic-solid coupling method. Thirdly, the dynamic response analysis under different parameters.

4.1. Modal and Meshing. Ansys Workbench integrates a variety of modules to facilitate the calculation and coupling analysis of flow and solid structures, hence numerical simulations of sea areas are conducted in Fluent module, while Modal, Modal Acoustics, Static Structural, Transient Structural, and other modules were used to analyze the dynamic characteristics and responses of structures.

As the tunnel pipe body and pier body are consolidated as a whole, and the both sides of the midspan are symmetrical, then the following part is taken as the research object, that is, "half span + one span + half span." In this simplified model, the dynamic response analysis is carried out on the stress condition of pier bottom and the displacement condition of pipe top, ignoring the internal interaction between pipe body and pier body. The form of the typical structure is shown in Figure 3(a).

The cross section width of the pipe body is D assumed, and the single-span length is L . Making the flow field fully developed and not affecting navigation, the basic size parameters of calculated water area are set as follows: the distance between the inlet and the upstream face is $10D$, the distance between the outlet and the lee face is $20D$, the water surface is placed 30 m above the top of the tunnel, the pier is 30 m high, and the axial length along the pipe body is $2L$, whereas the circular section is one of the most common and typical forms; a circular section with a diameter of D is selected as the tunnel pipe body.

As shown in Figure 3(a), the top and two sides of the water body were set as symmetry boundaries, the bottom of

the water body was set as no-slip wall, and the tunnel pipe surface and pier side surface (i.e., the fluid-solid interface) were set as no-slip wall. Symmetry constraints are set at both ends of the pipe body (remote displacement), and fix support is set at the bottom of the pier.

For the water area, SSTk- ω model is adopted considering the computational efficiency and the applicability of the model. Moreover, an encrypted boundary layer grid is set on the wall, as shown in Figure 3(b).

4.2. Hydrodynamic Coefficient Calculation. Reynolds number is a dimensionless parameter used to characterize fluid flow and reflects the ratio of inertia force and viscosity force. The calculation formula is as follows:

$$Re = \frac{U D}{\nu}, \quad (7)$$

where ν is the kinematic viscosity of the fluid, and the value of seawater is $9.738 \times 10^{-7} \text{ m}^2/\text{s}$ at 20°C .

Strouhal number is a parameter used to characterize the lift frequency (or vortex discharge frequency) and is related to Reynolds number and junction interface shape. Its definition can be written as follows:

$$St = \frac{f_s V}{D}. \quad (8)$$

The drag coefficient and lift coefficient in Fluent are calculated by dividing the lift and drag forces by the reference values of the dynamic pressures. The calculation formulas of drag and lift coefficients are as follows:

$$C_D = \frac{F_D}{0.5\rho_w A u^2}, \quad (9)$$

$$C_L = \frac{F_L}{0.5\rho_w A u^2},$$

where A is the area facing of water surface.

The average pressure coefficient calculated in Fluent is

$$C_p = \frac{p - p_\infty}{0.5\rho_w u^2}. \quad (10)$$

5. Results and Analysis

5.1. Analysis of Flow Field around Sections. Taking the actual flow field of a certain sea area as the research background, natural condition parameters are as follows: the density of seawater is about 1030 kg/m^3 , the viscosity coefficient is $0.001003 \text{ Pa}\cdot\text{s}$, the local coastal surface current velocity is about $0.6\text{--}0.8 \text{ m/s}$, and the maximum velocity can reach $2.55\text{--}3.06 \text{ m/s}$, therefore the selected flow rate ranges from 0.5 m/s to 4 m/s in order to make the research more practical, while the Reynolds number is between 5×10^6 and 4×10^7 .

The effects of different ocean current velocities on the flow characteristics have been investigated, focusing on the circumferential pressure and lift-drag correlation coefficients of the circular section of the pipe body.

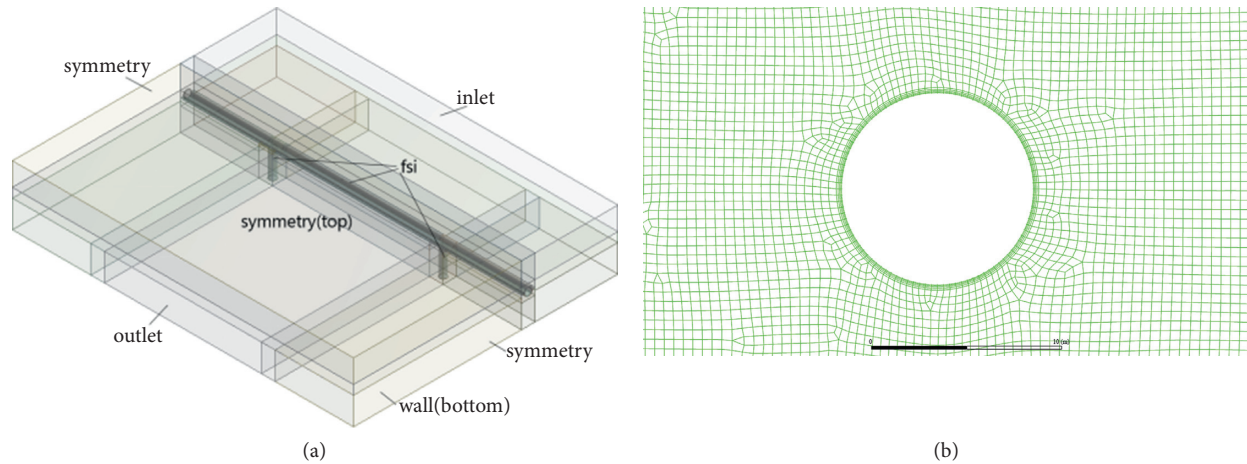


FIGURE 3: Computing domains and meshing. (a) Compute domain and boundary settings. (b) Local magnification of boundary layer grid.

5.1.1. Pressure Field Characteristic. As Figure 4 shows, with the increase of flow velocity, the circumferential pressure around the section increases sharply, and the pressure differential resistance between the front face and back face increases significantly. The pressure distribution curves under different flow rates show the same shape and trend, and the intervals of positive pressure and negative pressure are rarely changed.

The meaning of θ is illustrated in Figure 5. The maximum negative pressure is reached near the upper and lower vertexes of the cross section, that is, $|\theta|$ is nearly $90\sim 100^\circ$, and the absolute value of the maximum negative pressure is about 2.5~2.7 times of the maximum positive pressure. While $|\theta|$ is in the range of $0\sim 40^\circ$ on the back flow side, the negative pressure is low and basically unchanged, accessing the stability zone.

The distribution law of steady pressure around crossing section of tunnel body is consistent at different incoming flow rates. As shown in Figure 6, the distribution curves of the circumferential steady-state pressure coefficients at different flow rates almost coincide. With the velocity increasing, the flow separation point moves slightly closer to the back flow side. The absolute value of the ratio of the maximum negative pressure to the maximum positive pressure increases slightly as the flow rate increases.

5.1.2. Lift and Drag Coefficient Analysis. The time-history curves of lift coefficients and drag coefficients of the crossing section are shown in Figure 7.

The lift power spectrum can be obtained by FFT transformation of the time-history data of the lifting coefficient, and the lift power spectrum at different flow rates is shown in Figure 8.

It can be seen from Figures 7 and 8 that as the flow velocity increases from 0.5 m/s to 4 m/s, the vortex discharge frequency around the cross section gradually increases.

According to the time discrete data of lift and drag coefficients calculated in Fluent, the mean value and the variance value of these data were solved corresponding to steady-state coefficients and fluctuating coefficients.

Table 1 lists the results of the correlation coefficients about lift and drag around cross sections at different flow rates. Combined with the above analysis, it can be seen that the circumferential pressure of the section increases significantly in practice with the increase of the flow velocity, and the lift force and resistance of the structure also increase as well. However, as shown in Table 1, both the fluctuating lift coefficient and the steady-state drag coefficient decrease with the increase of velocity, while the values of steady-state lift coefficient and the fluctuating drag coefficient are small and fluctuate as the velocity changes. In addition, with the increase of flow velocity, the vortex discharge period gradually decreases, and the calculated Strouhal number has a small fluctuation within the process of velocity variation, and the Strouhal number of circular section fluctuates from 0.34 to 0.39.

5.2. Analysis of Structural Dynamic Characteristics. The water depth, velocity, and the size of the structure all affect the natural vibration characteristics, and appropriate structural parameters will reduce the possibility of vortex-induced resonance. Referring to the design of section of suspension tunnel [14, 15], the diameter of the circular section is about 10~25 m, while the length of a single span is about 100~500 m, hence the size of typical structure of underwater tunnels will be explained in the following sections.

5.2.1. Comparison of Dry and Wet Modes. In the basic example of wet modal analysis, pipe diameter D is 20 m, the distance between the two support piers L is 200 m, the pier diameter is 16 m, the pier is 30 m high, and the top of the tunnel pipe is placed 30 m underwater.

Dry and wet modals are simulated and conducted, respectively, and the natural vibration frequencies of each mode are listed in Table 2.

Through comparative analysis of dry and wet modes, it can be found from Table 2 that the natural vibration frequencies of wet modes are all smaller than the

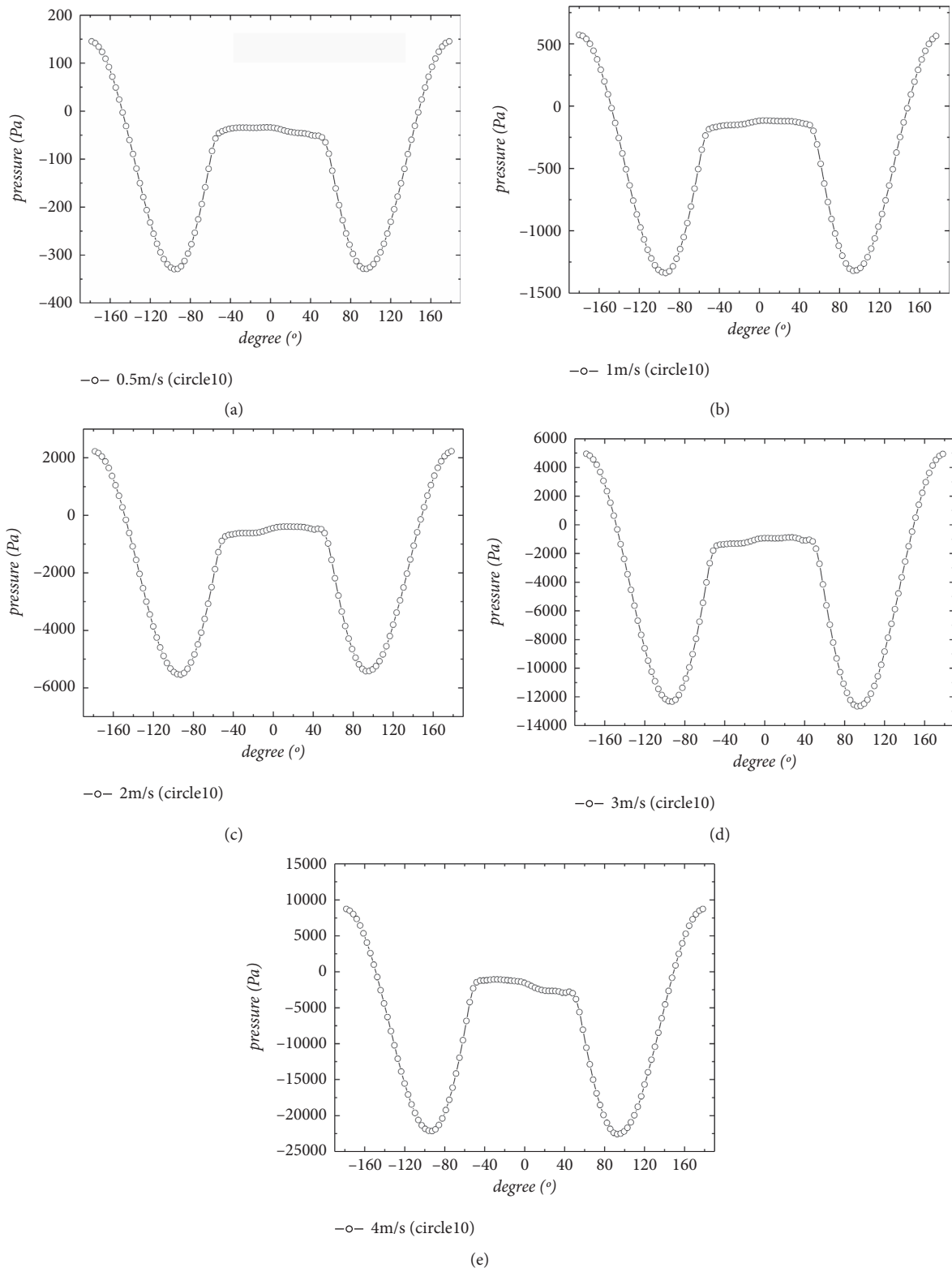


FIGURE 4: Steady circumferential pressure distribution at different velocities: (a) 0.5 m/s; (b) 1 m/s; (c) 2 m/s; (d) 3 m/s; (e) 4 m/s.

corresponding values of the structure in air. According to the basic idea of additional mass method, water acts as an additional mass applied to the structure. From the dynamic equation of the structure, it is evident that the

greater the additional mass, the lower the natural frequency of the structure, which demonstrates that the structure of numerical simulation is consistent with the theoretical practice.

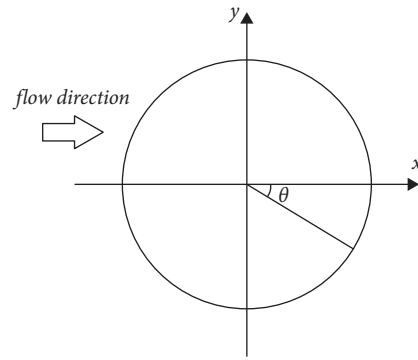


FIGURE 5: The meaning of θ in circumferential pressure diagram.

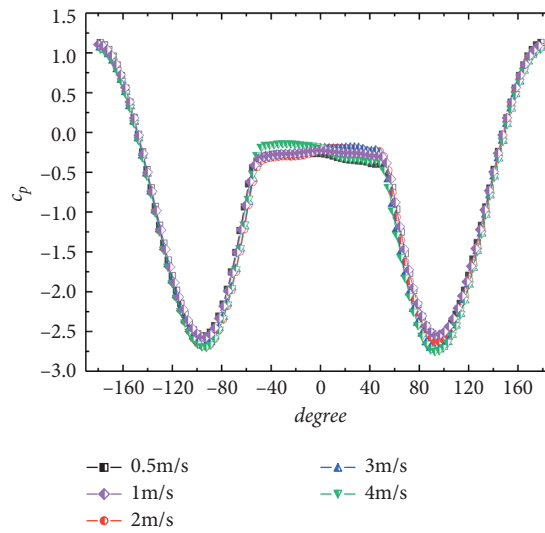


FIGURE 6: Steady circumferential pressure coefficient at different velocities.

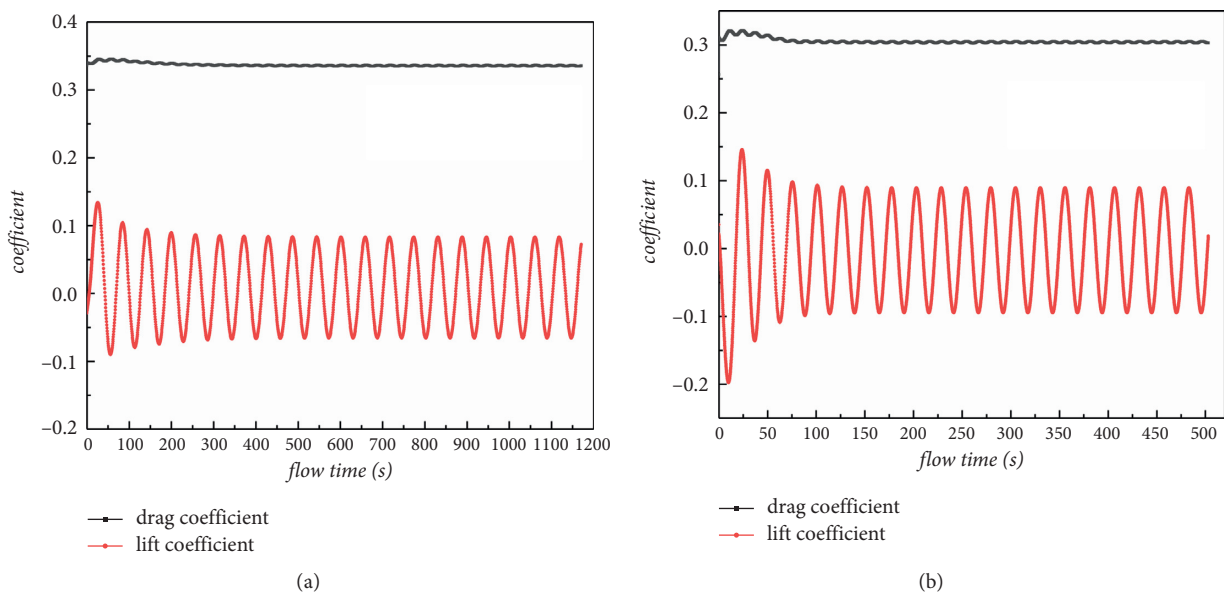


FIGURE 7: Continued.

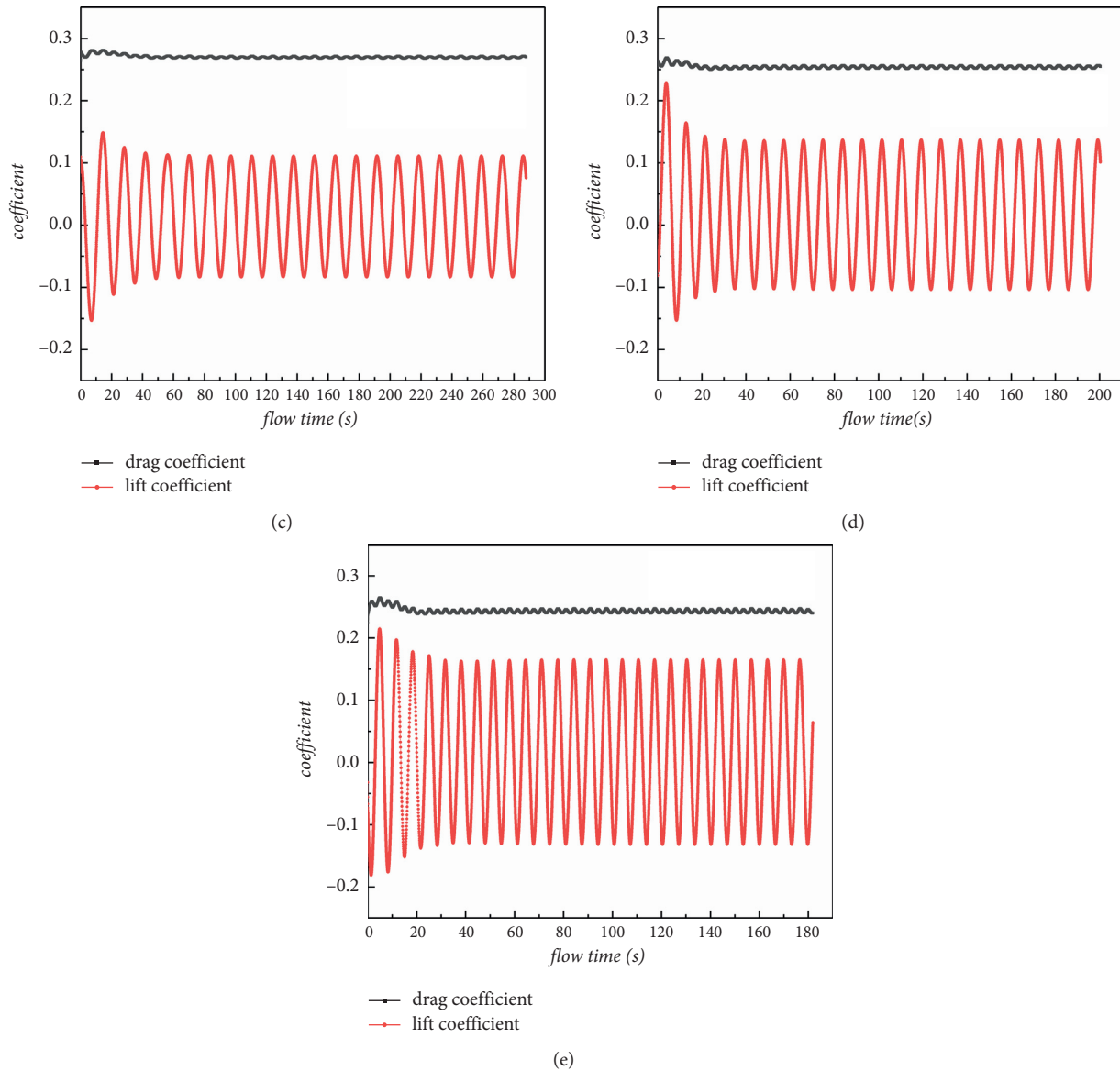


FIGURE 7: Time-history curves of lift coefficient and drag coefficient at different flow rates: (a) 0.5 m/s; (b) 1 m/s; (c) 2 m/s; (d) 3 m/s; (e) 4 m/s.

5.2.2. Effect of Placement Depth on Natural Vibration Characteristics. The range of the placement depth under water, that is, the distance between the top of the tunnel pipe and the water free surface was selected from 20 m to 150 m to explore the variation rule of the natural vibration frequency of the structure with the water depth.

The results of the first ten natural frequencies at different placement depths based on acoustic-structure coupling method are shown in Table 3.

Under the same water depth, the natural vibration frequency of the structure increases gradually as the modal order is higher. On the whole, although the natural vibration frequency of the structure decreases with the increase of water depth, the frequency drop is very little whose amplitude of variation is basically lower than 0.01 Hz per 10 meters of water depth.

It is proved that water depth is not the key factor affecting the natural vibration of the structure, but considering the environmental factors and construction conditions,

water depth is still an indispensable factor. When the water depth is shallow, it may cause disturbance to navigation and other aquatic production activities. When the water depth is too deep, it will bring great difficulties to construction and increase the cost of pipe transportation. In addition, the environment of high water pressure will reduce the stability of the structure. Therefore, the depth of placement should be determined after a comprehensive evaluation of economy and safety based on the actual sea conditions.

5.2.3. Effect of Width-Span Ratio on Natural Vibration Characteristics. The single-span was set to 100 m~450 m, and three sections with diameters of 20 m, 15 m, and 10 m were adopted, while the corresponding pier body was 16 m, 12 m, and 8 m in diameter. The top of tunnel pipe body is 30 m underwater and the pier is 30 m high. Figure 9 and Table 4 respectively show the classification of each case and

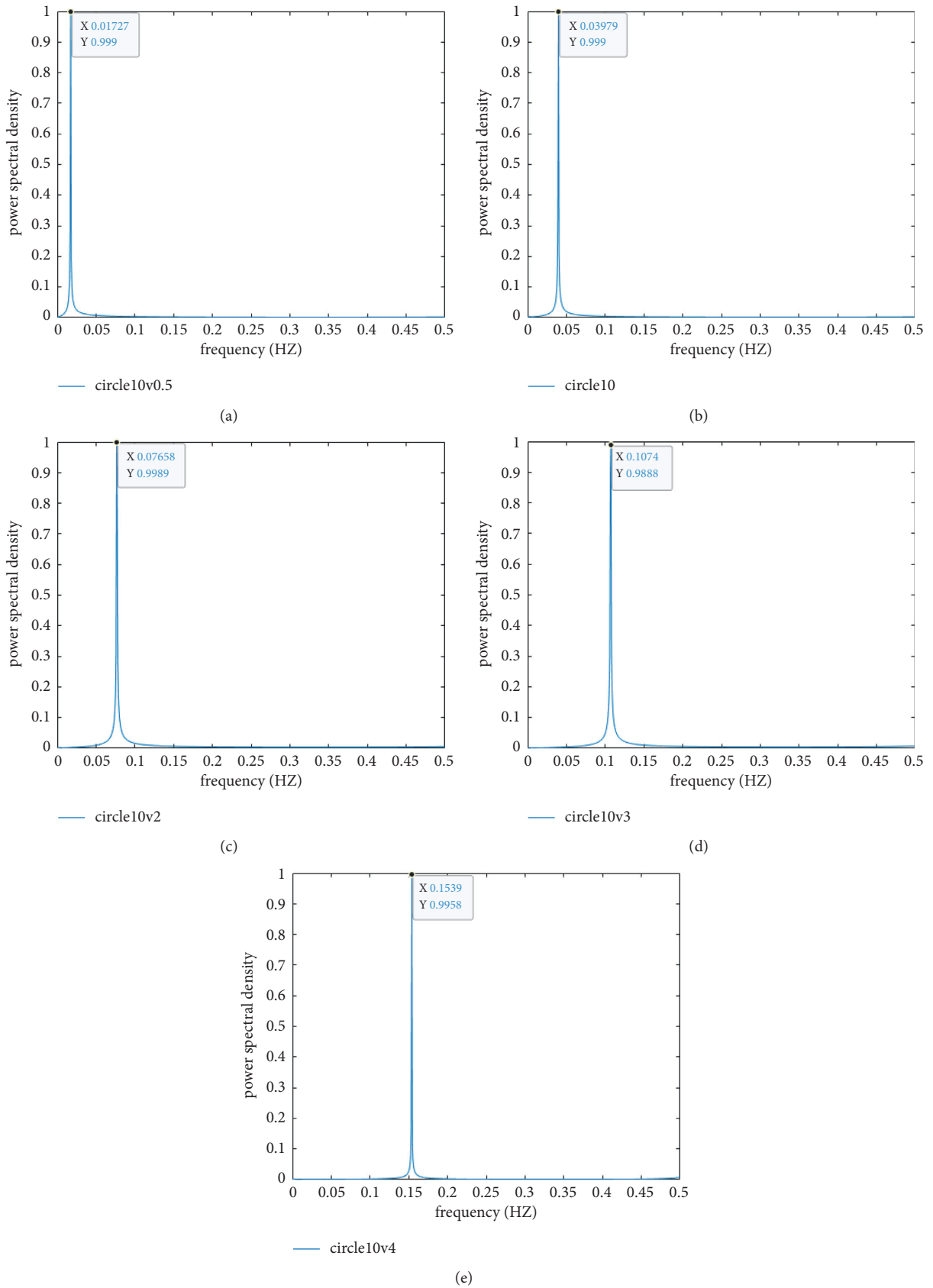


FIGURE 8: Lift power spectrum at different flow rates: (a) 0.5 m/s; (b) 1 m/s; (c) 2 m/s; (d) 3 m/s; (e) 4 m/s.

TABLE 1: Lift and drag coefficients at different flow rates.

Velocity (m/s)	Steady-state lift coefficient	Fluctuating lift coefficient	Steady-state drag coefficient	Fluctuating drag coefficient	Vortex period (s)	Strouhal number
0.5	0.012	0.241	0.358	0.009	58.11	0.3454
1	0.015	0.225	0.296	0.011	25.13	0.3979
2	0.012	0.193	0.278	0.008	13.07	0.3829
3	0.014	0.187	0.255	0.012	9.31	0.358
4	0.009	0.172	0.236	0.015	6.50	0.3845

TABLE 2: Natural frequencies of dry and wet modals.

Order	Dry modal frequencies, f_d/Hz	Wet modal frequencies, f_d/Hz	$(f_d - f_a)/f_d$
1	1.4468	0.66516	0.540
2	1.5081	0.71504	0.526
3	1.7819	0.82936	0.535
4	2.6011	1.2089	0.535
5	3.5976	1.7013	0.527
6	4.0424	2.2095	0.453
7	5.4398	2.1430	0.606
8	5.8327	2.6401	0.547
9	6.9219	2.8184	0.593
10	6.9655	3.4567	0.504

TABLE 3: The first ten natural frequencies at different water depths (Hz).

Order	Depth of tunnel pipe roof (m)							
	20	30	40	50	60	80	100	150
1	0.66584	0.66516	0.66263	0.66075	0.65984	0.65692	0.65356	0.64704
2	0.71406	0.71504	0.71321	0.71278	0.7125	0.71134	0.7095	0.70527
3	0.82964	0.82936	0.82833	0.82744	0.82767	0.82601	0.82398	0.81834
4	1.2091	1.2089	1.2077	1.207	1.2069	1.2054	1.2027	1.1956
5	1.7021	1.7013	1.70	1.6979	1.6962	1.693	1.6875	1.6758
6	2.21	2.2095	2.2073	2.2069	2.2073	2.2054	2.2039	2.1988
7	2.1458	2.143	2.1411	2.1387	2.1371	2.1351	2.1326	2.1294
8	2.6414	2.6401	2.6384	2.6369	2.6359	2.6326	2.6265	2.6115
9	2.8153	2.8184	2.8134	2.8116	2.8108	2.807	2.8015	2.7884
10	3.4578	3.4567	3.4526	3.4513	3.4525	3.4493	3.4456	3.4279

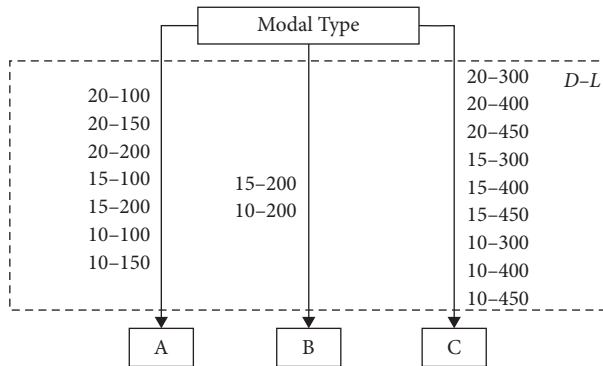


FIGURE 9: Vibration mode classification.

the range of width-span ratio with the corresponding description of vibration mode shape for each type.

According to the results of the first three vibration modes among 18 cases, it can be found that, from the perspective of mode shapes, with the span length gradually

TABLE 4: Modes with different width-span ratios.

Type	Width-span ratio	Mode shape
A	0.067~0.2	1 st Bending of pier body
		2 nd Horizontal bending of pipe
		3 rd Vertical bending of pipe
B	0.05~0.075	1 st Horizontal bending of pipe
		2 nd Bending of pier body
		3 rd Vertical bent of pipe
C	0.022~0.05	1 st Horizontal bending of pipe
		2 nd Vertical bending of pipe
		3 rd Second-order horizontal bending of pipe

increasing, the rule of mode has also changed. As the width-span ratio reaches a certain range, the vibration modes can be divided into three types. Type A occurs when the single-span length is short, while the width-span ratio ranges from 0.067 to 0.2, the first-order vibration mode is mainly the bending vibration of pier body with the bottom of pipe

totally fixed, while the tunnel pipe body vibrates back and forth in the horizontal direction as a whole. The 2nd and 3rd mode shapes of type A are respectively the horizontal bending vibration and vertical bending of pipe. Then comes the type B as the single-span increases to about 200 m, while the width-span ratio fluctuates between 0.05 and 0.075. Type B is almost like a transitional form presented and the first-order vibration mode turns to be the first-order bending of the pipe body in the horizontal direction, while the 2nd vibration mode is the bending of the pier with the pipe body vibrating back and forth horizontally as a whole, and the 3rd mode shape is the same as Type A. When the span length increases further, the width-span ratio is within a small range from 0.022 to 0.05, the first-order vibration mode is mainly the first-order horizontal bending of the tunnel pipe, with piers torsion, and the 2nd mode shape is vertical bending vibration of pipe, whereas the 3rd vibration mode shape is totally different, where the second-order horizontal bending of pipe body comes up.

From the perspective of natural frequencies, the frequency results of 18 cases are listed in Tables 5–7. With same span length, the natural frequencies increase as modal order goes up. Within the variation range of single-span length from 100 m to 450 m, the width-span ratio of 20 m pipe body varies from 0.2 to 0.22, while the corresponding ratio, for a pipe with a diameter of 15 m and 10 m, ranges respectively from 0.15 to 0.033 and from 0.10 to 0.022.

As is shown from Tables 5–7 and Figure 10 that the natural frequency decreases gradually with the increase of the span length, and the rate of the decrease gradually slows down, when the single-span length increases to 300 m, the first three frequencies of the structure basically tend to coincide, with a small increase per each mode.

According to the modal order, the variation diagram of natural vibration frequency with width-span ratio was drawn at the same order, as shown in Figure 11. It can be found that there is a strong positive linear relationship between the natural frequency and the width-span ratio of the typical structure, and the vibration frequency increases gradually as the ratio increases in each mode. Besides, under the same width-span ratio, the natural vibration frequency decreases with the increase of pipe diameter; however, while under the same span length, the natural vibration frequency of the structure increases with the increase of pipe diameter, which fully indicates that the span length has a stronger influence than the cross section size of tunnel pile body on the natural vibration frequency of the assumed structure.

It is a safer design to reduce the span length and avoid the fatigue or damage of the structure due to vortex-induced vibration resonance, which can also improve the stability of the structure. Nevertheless, surplus supporting piers can lead to a substantial rise in the cost of construction. Therefore, the proper structure size should be selected according to the actual water area characteristics and the principle of safety and economy.

5.3. Dynamic Response Analysis

5.3.1. Effects of Different Flow Rates. In the studies of the dynamic response of the structure at different flow rates, a circular section with a diameter of 10 m ($D=10$ m) was

selected for the pipe cross section, and the single-span length between the two piers was 200 m ($L=200$ m), while the calculated water area was $400\text{ m} \times 300\text{ m} \times 60\text{ m}$ according to the set above. The study subjects about dynamic response of the structure mainly include the displacement response of the pipe top in all directions and the bending moment as well as the shear force at the bottom of the pier; moreover, the stress, at the junction between the pipe body and the pier column, is also an important object that needs to be monitored. Table 8 shows the dynamic response results of the structure at different flow rates in the range from 0.5 m/s to 4 m/s.

In Table 8, Z_{\max} and Y_{\max} are the maximum horizontal displacement and the maximum vertical displacement of pipe top, Moment_{\max} and Shear_{\max} are the maximum bending moment and the maximum shear of pier bottom, and Stress_{\max} is the maximum stress at the junction between the pipe and pier.

The main bending moment at the bottom of the pier is axial around the pipe body, showing that the bending of pier bodies occur in the plane where the water cross section is located (YZ plane). As for the shear force of pier bottom mainly acts on the section along the counter-current direction (Z -axis negative direction). The numerical results of the maximum bending moment and the maximum shear force at the bottom of the pier, as well as the maximum stress at the junction of the pipe body and the pier body changing with the flow rate, are plotted in Figure 12. As is depicted that with the increase of the flow rate, the bending moment and shear force at the bottom of the pier gradually increased, so as the stress at the junction, and the increase amplitude was nonlinear, showing a quadratic increasing trend.

Under the action of ocean current load, the structure displacements occurred in both the downstream direction and the transverse direction, and the maximum displacements occur at the top of the tunnel pipe section, thus the top part of the pipe was set as the monitored path. Figure 13 shows the vertical and horizontal displacement curves of the path along the longitudinal axis. It can be seen from the figure that the displacement of the pipe body is larger along the ocean current direction than the displacement in the cross current direction. In the vertical direction, the displacements of adjacent spans are in the opposite direction, and the maximum displacement is achieved at the middle of the span, while there is nearly no vertical displacement at the pier support. In the horizontal direction, the horizontal displacement of adjacent spans is in the same direction, and the maximum displacement is achieved at the middle of the span as well, while the minimum displacement is achieved at the pier support. With the increase of flow velocity, the displacement at each point along the pipe also gradually increases, with presenting a quadratic growth trend.

5.3.2. Effects of Different Span Lengths. A circular section with a diameter of 10 m ($D=10$ m) is selected in the tunnel pipe section, and the single-span length L between the two piers is 100 m, 200 m, 300 m, 400 m, and 450 m. Table 9

TABLE 5: Natural vibration frequencies of structures with different span lengths ($D = 20$ m).

Diameter (m)	Span length (m)	Width-span ratio	Frequency (Hz)					
			1 st	2 nd	3 rd	4 th	5 th	6 th
20	100	0.200	1.1554	2.5646	2.8787	3.3667	4.0364	5.6344
	150	0.133	0.90607	1.2315	1.4076	1.9517	2.5616	2.972
	200	0.100	0.66682	0.71533	0.8294	1.209	1.7053	2.2106
	300	0.067	0.34852	0.40492	0.46432	0.57254	1.1229	1.2915
	400	0.050	0.2076	0.24639	0.29847	0.32926	0.70337	0.76766
	450	0.044	0.16769	0.19682	0.24235	0.26155	0.57159	0.61213

TABLE 6: Natural vibration frequencies of structures with different span lengths ($D = 15$ m).

Diameter (m)	Span length (m)	Width-span ratio	Frequency (Hz)					
			1 st	2 nd	3 rd	4 th	5 th	6 th
15	100	0.15	0.92597	1.9591	2.2269	3.0122	3.072	5.1569
	150	0.10	0.73692	0.9308	1.0998	1.6272	1.898	2.9244
	200	0.075	0.55022	0.59316	0.67017	0.96852	1.3962	1.8702
	300	0.050	0.26181	0.31893	0.36445	0.44276	0.85128	1.0218
	400	0.0375	0.1553	0.18946	0.22918	0.25303	0.5329	0.5951
	450	0.033	0.12601	0.15151	0.18604	0.20011	0.4363	0.47561

TABLE 7: Natural vibration frequencies of structures with different span lengths ($D = 0$ m).

Diameter (m)	Span length (m)	Width-span ratio	Frequency (Hz)					
			1 st	2 nd	3 rd	4 th	5 th	6 th
10	100	0.100	0.73518	1.3421	1.5647	2.1748	2.4161	4.2544
	150	0.067	0.57205	0.63435	0.77965	1.1656	1.3685	2.1223
	200	0.05	0.36461	0.3882	0.44204	0.66155	0.8898	1.2757
	300	0.033	0.17839	0.223	0.25877	0.3004	0.59778	0.70932
	400	0.025	0.10585	0.12946	0.15788	0.1691	0.36622	0.40606
	450	0.022	0.083829	0.10547	0.12612	0.13562	0.28867	0.32777

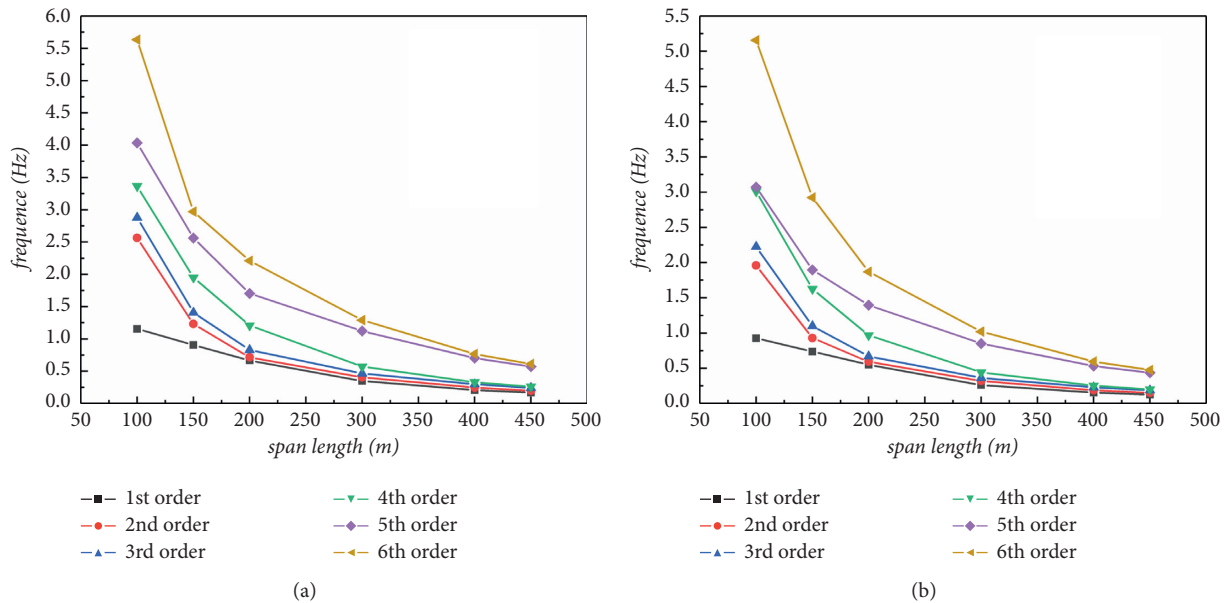


FIGURE 10: Continued.

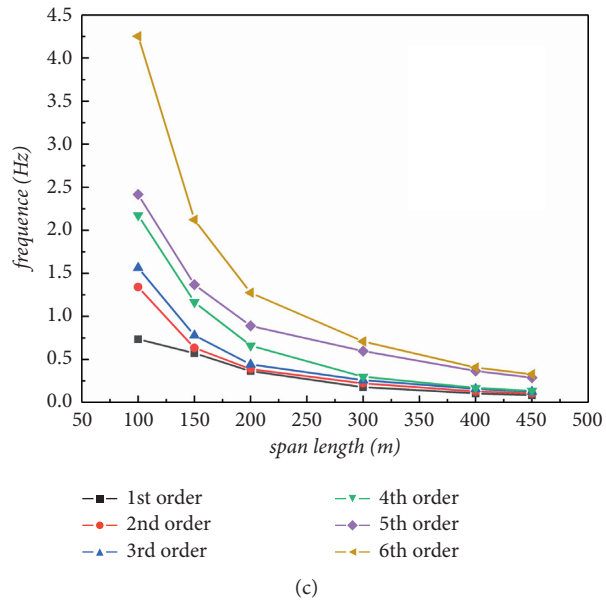
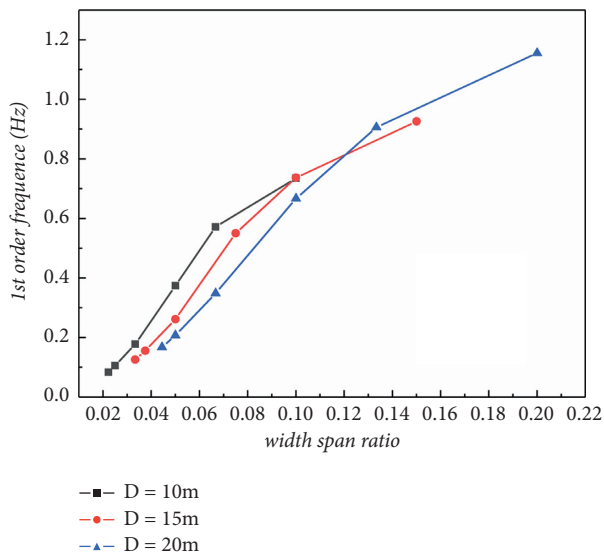
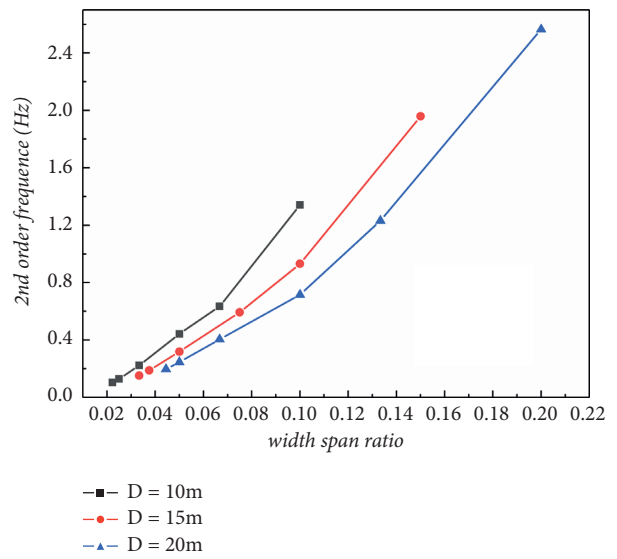


FIGURE 10: Line chart of natural vibration frequency-span with different pipe diameters: (a) $D = 20$ m; (b) $D = 15$ m; (c) $D = 10$ m.



(a)



(b)

FIGURE 11: Continued.

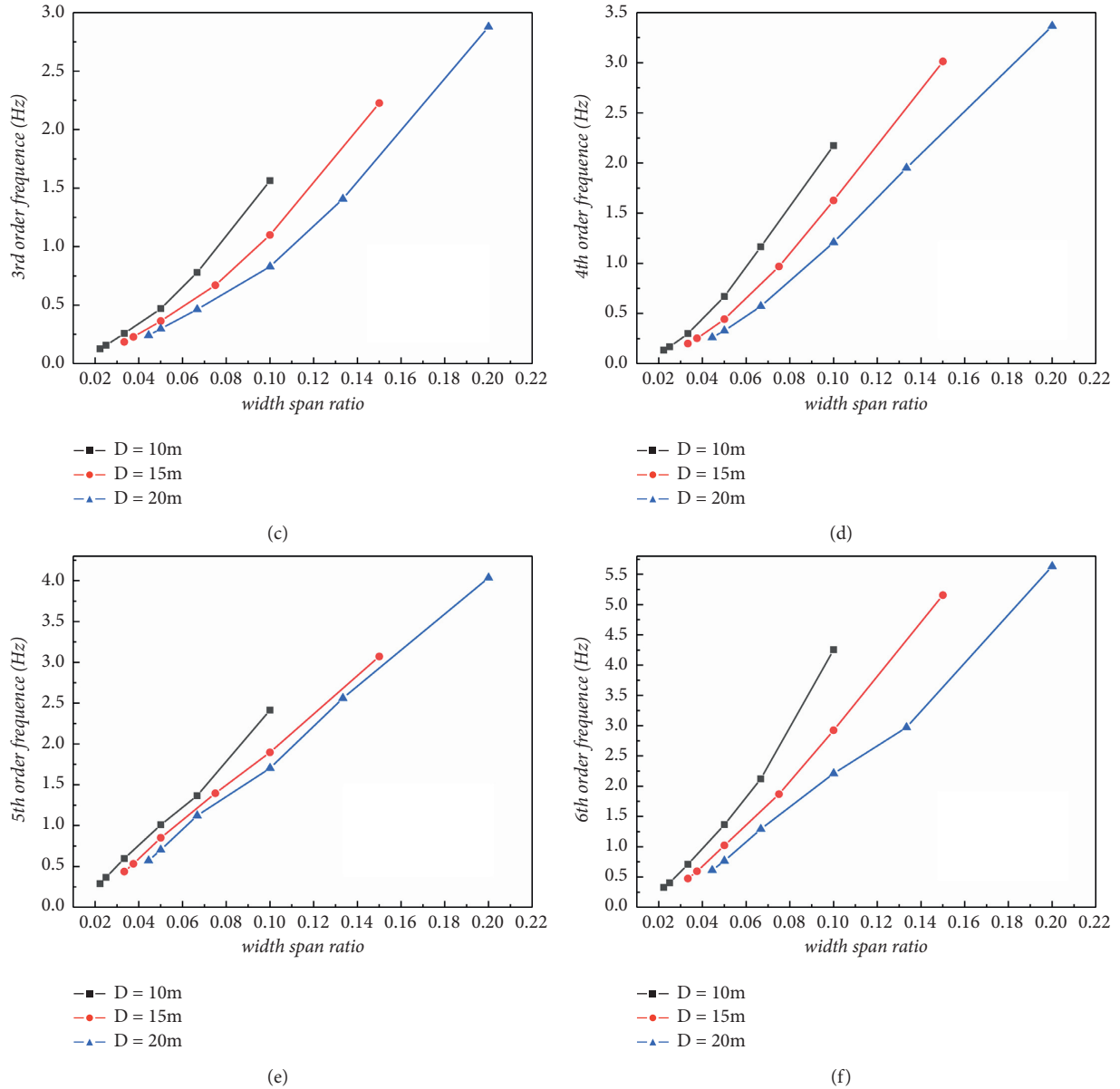


FIGURE 11: Line chart of natural frequency-width-span ratio in each mode: (a) 1st order; (b) 2nd order; (c) 3rd order; (d) 4th order; (e) 5th order; (f) 6th order.

TABLE 8: Dynamic response results of structures at different flow rates.

Velocity (m/s)	0.5	1	2	3	4
Z_{\max} (mm)	1.258	4.979	19.91	44.62	79.03
Y_{\max} (mm)	0.415	1.214	4.549	10.10	26.69
Moment _{max} (MN·m)	18.66	73.617	292.24	651.34	1152.1
Shear _{max} (MN)	0.569	2.241	8.930	19.918	35.523
Stress _{max} (MPa)	0.165	0.632	2.615	5.827	10.306

shows the dynamic response results of the structure within the span range of 100~450 m.

Figure 14 shows the numerical results of the maximum bending moment and shear force at the bottom of the pier, as well as the stress at the junction between the pipe body and the pier with the single-span length changing. It can be

found that the maximum of the bending moment and shear force at the bottom of the pier gradually increase with the span length increasing, which shows a linear growth trend.

The tunnel pipe top along the axis was set as the monitoring path; Figure 15 shows the axial distribution curves of vertical and horizontal displacements of pipe top

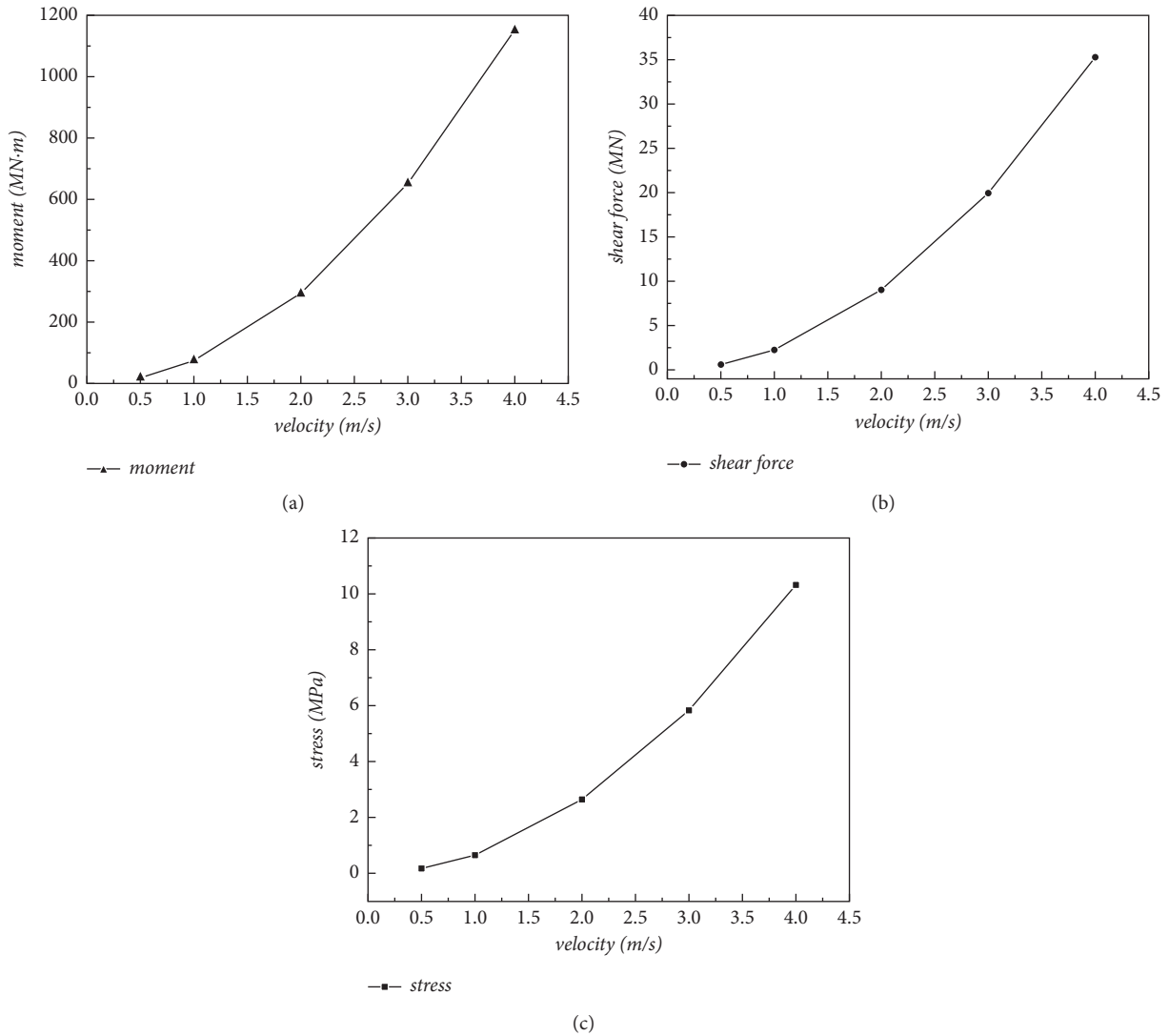


FIGURE 12: Line chart of dynamic response at different flow rates. (a) Maximum moment of pier bottom $Moment_{max}$. (b) Maximum shear of pier bottom $Shear_{max}$. (c) Maximum stress at the junction $Stress_{max}$.

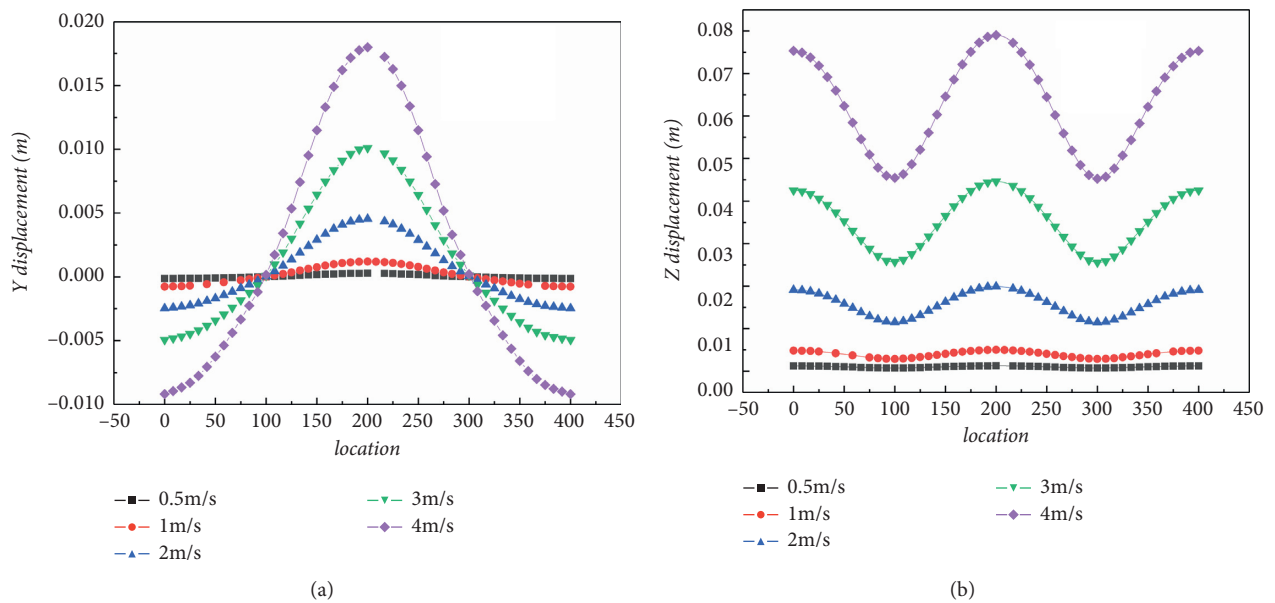
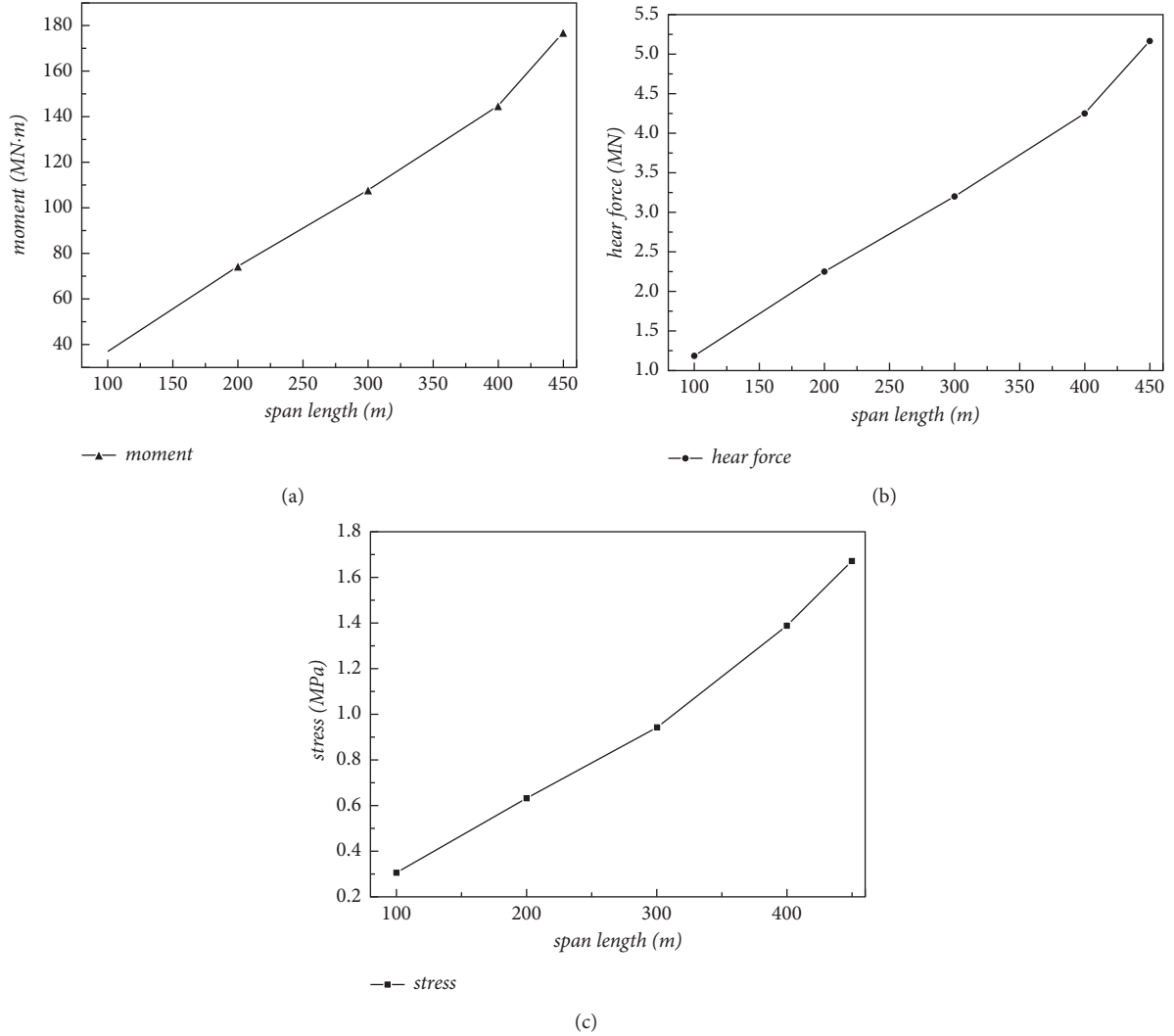


FIGURE 13: Axial displacement of the top of the pipe. (a) Vertical displacement. (b) Horizontal displacement.

TABLE 9: Dynamic response results of structures with different single-span lengths.

Span (m)	100	200	300	400	450
Z_{\max} (mm)	1.626	4.979	17.146	42.004	70.667
Y_{\max} (mm)	0.095	1.214	1.741	2.059	2.12
Moment _{max} (MN·m)	37.14	73.62	107.45	144.20	176.32
Shear _{max} (MN)	1.187	2.241	3.205	4.249	5.163
Stress _{max} (MPa)	0.302	0.632	0.942	1.388	1.671

FIGURE 14: Line chart of dynamic response under different spans. (a) Maximum moment of pier bottom Moment_{max}. (b) Maximum shear of pier bottom Shear_{max}. (c) Maximum stress at the junction Stress_{max}.

under different single-span lengths. It can be found that, in the horizontal direction (Z direction), adjacent pipe segments still maintain the same displacement along the flow direction, and the maximum displacement occurs at the middle of the span, while the minimum displacement is located at the pier top support. With the increase of the span length, the horizontal displacement at each point gradually increases.

In the vertical direction (Y direction), if the span exceeds a certain range, the displacement response will

change. When the single-span length is less than 300 m, the vertical displacements of adjacent spans are reversed, while when the single-span length is longer than 300 m, the vertical displacements of adjacent spans are in the same direction, and the maximum displacements are achieved at the middle of the span, while the vertical displacements at the pier top support are almost zero. With the increase of span length, the maximum of vertical displacement increases gradually.

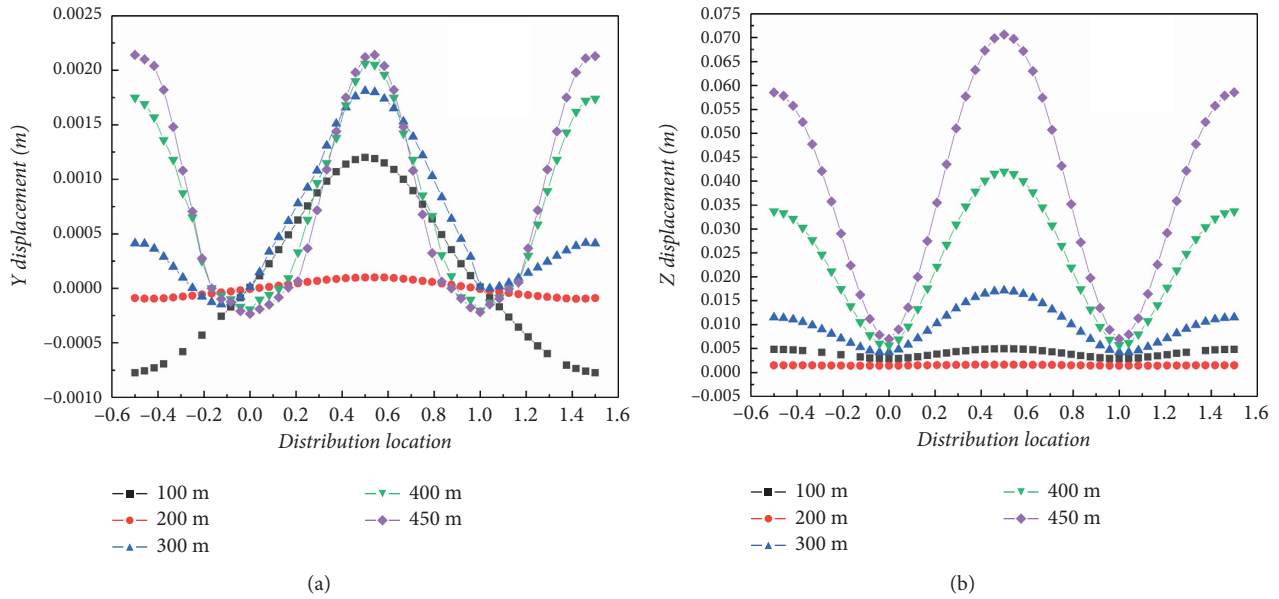


FIGURE 15: Axial displacement of the top of the pipe. (a) Vertical displacement. (b) Horizontal displacement.

6. Conclusion

A new underwater traffic structure, named as deep water highway tunnel (DWHT), is proposed by comprehensively comparing existing different traffic modes across broad waters. DWHT is a new and competitive solution to the challenge of crossing wide water areas.

In this paper, a simplified numerical simulation model of DWHT (half span + one span + half span) is proposed firstly. Then, the characteristics of flow field around the circular cross section of the underwater tunnel are studied. In addition, dynamic characteristics and dynamic response of the typical structure of DWHT under the action of ocean current are explored. It is found that ocean current velocity and single-span length of tunnel body are the most important influencing factors. Specific conclusions are as follows:

- (1) Although the variation of flow velocity has little effect on the distribution of circumferential pressure around the circular cross section, the increase of flow velocity can significantly increase the circumferential pressure, as well as the lift force and drag force. However, the fluctuating lift coefficient and steady-state drag coefficient decrease as velocity increases. At the same time, with the increase of the flow velocity, the vortex discharge period gradually decreases, and the Strouhal number fluctuates within a relatively stable range. The Strouhal number of the circular section fluctuates between 0.34 and 0.39.
- (2) Water can increase the additional mass and reduce the natural vibration frequency of the structure in wet modal. The water depth has unconsidered effect on the natural vibration frequency while the tunnel structure is completely submerged in water, and the natural vibration frequency decreases slightly with the increase of water depth.

- (3) When the pipe span exceeds a certain range, the vibration mode shape at the same order will change. With the increase of single-span length, the natural vibration frequency of the structure decreases gradually. While the span length exceeds 200 m (the corresponding diameter is 10–20 m), the first-order vibration mode changes from pier bending to pipe horizontal bending.
- (4) The dynamic responses of the structure have shown the quadratic increasing trend with the increase of flow velocity, including the bending moment, the shear at pier bottom, and the stress at the junction between the pipe body and the pier, as well as the displacements. Besides, the dynamic response results increase linearly with the span length.
- (5) The maximum displacement of the structure is located at the middle of the top span of the pipe. When the single-span length is short, the vertical displacement of adjacent spans is opposite and the horizontal displacement is in the same direction. When the single-span length is relatively longer, the vertical and horizontal displacements of adjacent spans are both in the same direction.

There still are some limitations in this research, for instance, the internal force between pier body and tube body is not considered, or the variation characteristics of three-dimensional flow field along the axial direction of the tube body are not emphasized, and the quantitative law of the influence of span length (or width-span ratio) on the natural vibration characteristics needs to be further explored.

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 to report regarding the present study.

References

- [1] R. Svein, J. Leira Bernt, K. M. Okstad, K. M. Mathisen, and T. Haukås, “Dynamic response and fluid structure interaction of submerged floating tunnels,” *Computers & Structures*, vol. 72, no. 4, 1999.
- [2] H. Kunisu, T. Fujii, Y. Mizuno, and H. Saeki, “The study of submerged floating tunnel characteristic under the wave conditions,” in *Proceedings of the 4th International Offshore and Polar Engineering Conference*, vol. 8, Osaka, Japan, April 1994.
- [3] M. Di Pilato, F. Perotti, and P. Fogazzi, “3D dynamic response of submerged floating tunnels under seismic and hydrodynamic excitation,” *Engineering Structures*, vol. 30, no. 1, 2007.
- [4] S. Tariverdilo, J. Mirzapour, M. Shahmardani, R. Shabani, and C. Gheyretmand, “Vibration of submerged floating tunnels due to moving loads,” *Applied Mathematical Modelling*, vol. 35, no. 11, 2011.
- [5] S. I. Seo, M. Hyung-suk, L. Jin-ho, and K. Jin-ha, “Simplified analysis for estimation of the behavior of a submerged floating tunnel in waves and experimental verification,” *Marine Structures*, vol. 44, 2015.
- [6] J. Mai, Z. Luo, and B. Guan, “The vortex-excited dynamic response for a submerged floating tunnel under the combined wave and current effect,” *Journal of Southwest Jiaotong University*, no. 5, pp. 600–604, 2004, (in Chinese).
- [7] G. Wang, X. Zhou, and B. Gao, “Flow Resistance analysis of submerged floating tunnel,” *Journal of Southwest Jiaotong University*, vol. 178, no. 6, pp. 715–719, 2007, (in Chinese).
- [8] G. Luo, X. Zhou, C. Zhang, and D. F. Li, “Analysis on Characteristics of Flow Passing submerged floating tunnels of different sections,” *Journal of the China Railway Society*, vol. 35, no. 1, pp. 115–120, 2013, (in Chinese).
- [9] Q. Li, S. Jiang, and C. Xiang, “Experiment on pressure characteristics of submerged floating tunnel with different section types under wave condition,” *Polish Maritime Research*, vol. 25, no. s3, 2018.
- [10] Z. Fan and Y. Yuan, “Dynamic response analysis of submerged floating tunnel supported on columns in vortex-induced vibration,” *Journal of Railway Science and Engineering*, vol. 17, no. 3, pp. 653–659, 2020, (in Chinese).
- [11] Y. Yang, Y. Xiang, H. Lin, and Z. Chen, “Study on vibration response of submerged floating tunnel considering vehicle eccentric load,” *Applied Ocean Research*, vol. 110, Article ID 102598, 2021.
- [12] H. Zhang, Z. Yang, J. Li et al., “A global review for the hydrodynamic response investigation method of submerged floating tunnels,” *Ocean Engineering*, vol. 225, Article ID 108825, 2021.
- [13] A. Donna, “Chapter 10 Submerged floating tunnels— a concept whose time has arrived,” *Tunnelling and Underground Space Technology Incorporating Trenchless Technology Research*, vol. 12, no. 2, 1997.
- [14] X. Chen, Z. Chen, S. Cai et al., “Numerical investigation of dynamic responses and mooring forces of submerged floating tunnel driven by surface waves,” *Scientific Reports*, vol. 10, no. 1, Article ID 18836, 2020.
- [15] B. Jiang, B. Liang, and S. Wu, “Feasibility study on the submerged floating tunnel in Qiongzhou strait, China,” *Polish Maritime Research*, vol. 25, no. s2, 2018.

Research Article

Spatial-Temporal Graph Convolutional Framework for Yoga Action Recognition and Grading

Shu Wang 

School of Physical Education, Inner Mongolia Minzu University, Tongliao, Inner Mongolia 028000, China

Correspondence should be addressed to Shu Wang; wangshu2012@imun.edu.cn

Received 20 January 2022; Revised 8 February 2022; Accepted 24 February 2022; Published 29 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Shu Wang. 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 rapid development of the Internet has changed our lives. Many people gradually like online video yoga teaching. However, yoga beginners cannot master the standard yoga poses just by learning through videos, and high yoga poses can bring great damage or even disability to the body if they are not standard. To address this problem, we propose a yoga action recognition and grading system based on spatial-temporal graph convolutional neural network. Firstly, we capture yoga movement data using a depth camera. Then we label the yoga exercise videos frame by frame using long short-term memory network; then we extract the skeletal joint point features sequentially using graph convolution; then we arrange each video frame from spatial-temporal dimension and correlate the joint points in each frame and neighboring frames with spatial-temporal information to obtain the connection between joints. Finally, the identified yoga movements are predicted and graded. Experiment proves that our method can accurately identify and classify yoga poses; it also can identify whether yoga poses are standard or not and give feedback to yogis in time to prevent body damage caused by nonstandard poses.

1. Introduction

Yoga has become a very trendy fitness exercise in today's life. But yoga is much more than just a fitness exercise. Yoga is a physical and mental discipline that combines art, science, and philosophy. Yoga can help people regulate their breathing, keep their bodies healthy, and also calm their moods. In today's highly developed Internet, according to incomplete statistics, yoga has become the preferred fitness exercise for 300 million people [1]. As a scientific exercise, yoga encompasses breath control exercises, body stretching exercises, and mind cleansing [2]. Yoga originally originated in ancient India, then spread to the West, where it became a mainstream Western fitness modality, and then eventually spread globally with the Internet, becoming one of the most popular exercise cultures worldwide [3]. According to a joint UK and US survey, the demographic profile of the yoga training population found in the demographics indicates that women are the main enthusiasts of the sport, accounting for 85% of the total number of yoga practitioners [4–6].

Numerous studies have proven that yoga exercises are beneficial to the human body. There is also a large amount of research in rehabilitation on how to make yoga training work better for patients in their recovery process. This is one of the reasons why yoga has become a favorite exercise for many people [7]. In addition, research has proven that yoga has a complementary healing effect in the direction of eating disorders; it can modify the patient's eating habits and keep diet [8]. In the interviews of yoga practitioners, it was learned that yoga gave them a positive and subjective life experience, making them healthier and living an optimistic life. There were significant improvements in self-care, self-activity, life comfort, and dwelling senses [9–11]. In fact, most of the experience that yoga brings to people comes from the yoga instructor. The instructor, as the guide of yoga, influences the yoga student in an invisible way with his or her philosophy of teaching, teaching environment, outlook on life, values, and demonstration of yoga effectiveness [12].

Although some researchers have demonstrated that yoga can be practiced without differentiating between

“traditional” and “authentic” issues [13], most people currently prefer modern yoga. Modern yoga is simpler and less demanding in terms of postural alignment and breathing exercises [14]. This is one of the reasons why modern yoga has turned into a healthy exercise for young and old alike. However, due to the overall economic development, yoga has gradually become commercialized. With the commercialization, the expression of yoga has become diversified and more and more people have become attracted to yoga. In our literature research, we found that yoga is becoming a synonym for young, beautiful, and hot women [15]. Yoga can be found in various fashion magazines and shows yoga poses that have a certain ornamental quality and at the same time these poses are difficult in the eyes of professionals. For ordinary people, they are more attracted by the ornamental poses of yoga, but these poses are risky for them. Commerce has made yoga idealized in order to facilitate promotion and thus attract consumers [16]. However, the commercialization of yoga is also a double-edged sword. Consumers are likely to cause irreversible damage to their bodies in the process of blindly imitating yoga poses due to the unknown nature of the poses, which is a potential risk in yoga training.

Traditionally, yoga is taught face-to-face, with the yoga instructor instructing in person whether the yoga poses are standard or not. This kind of teaching can make yoga students have a more direct feeling of standard yoga movements. However, with the advent of the 5G era and the rapid development of short videos, short video platform bloggers often adopt online teaching methods to teach yoga poses in order to attract fans. This is also the way most people learn yoga at present. Most people choose to watch videos while imitating to achieve the purpose of learning yoga. However, most people do not have professional yoga equipment and props, and they are not clear enough about the standard yoga postures. Blindly imitating the yoga postures in the videos has a great risk of physical injury. To solve this problem, in our work, we propose to use real-time posture detection technology to detect posture movements of yoga students and then use deep learning algorithms to grade and match yoga movements. A reference movement is given to the yoga students, and for the nonstandard movements, the yoga students are prompted in time to prevent the occurrence of physical injuries. In the specific experiment, we use the deep camera to capture the training postures of yoga students and decompose the postures to understand the yoga movements from the computer level. The postures are then compared with a standard database to verify whether the postures are standardized and to give feedback to the yoga students. Experiments show that the method proposed in our research can provide effective feedback to yoga trainees on the grading of yoga poses. The contributions of this paper can be summarized as follows.

The rest of this paper is organized in the following manner. Section 2 discusses the work related to deep camera and action recognition. Section 3 introduces the skeleton recognition principle of graph convolution, then introduces the residual unit and multistream input structure, and finally introduces the optimization principle of the partial perception framework. Section 4 reports the experimental data

collection, model training details, and analysis of experimental results. Finally, Section 5 concludes our research and reveals some further research work.

2. Related Work

The presentation of human motion postures in 3D space often requires the use of depth cameras. Information such as joint angles and skeletal space points can be deduced from the depth camera or the spatial position data of the human body [17]. Different poses can generate different skeletal contours, and to solve this problem, some researchers have proposed the idea of spatial segmentation, which takes an approximate mapping approach to define the location of spatial points for each segmented region. Literature [18] proposed a joint distribution method, which takes a bidirectional derivative approach to the mapping function. Literature [19] also uses the joint distribution rule, and unlike the former, the method adopts a Bayesian algorithm to obtain the image contour conditions. The final distribution of the image contour conditions will be mapped to the hybrid framework to obtain the spatial distribution features. Literature [17] additionally uses learning conditional distributions when learning features in the hybrid framework to obtain the image contour features more directly. In [20], to solve the image contour error problem caused by pose ambiguity, the researcher distributed three depth cameras into different angles to capture the human motion contour in all directions and obtained the skeletal spatial position from a standard dataset. In [21], the researcher used the SVM method to learn different pose features and perform pose prediction in the acquired 3D shape data. This proposed method links contours and 3D shapes but requires the support of large databases. For motion capture depth cameras, calibration of the depth camera is also required to ensure accuracy in 3D reconstruction work. In [22], the researcher applied the EM algorithm to calibrate the human action pose for multicamera linkage, and the mapping of 2D contours to 3D skeletal joints was achieved by training a neural network. Literature [23] adopts hybrid probabilistic PCA to predict the 3D body structure captured by the depth camera, which improves the 3D joint point coordinate accuracy.

Human motion recognition techniques originated from skeletal annotations [24, 25] by video clips [26, 27] to obtain the motion pose of each frame, which was then obtained by manual criteria. Previous human action recognition methods are based on RGB images, but this method is limited to the influence of nonobjective environments. The human skeleton-based action recognition method is less influenced by the nonobjective environment. This method can acquire the spatial-temporal features between joint points and learn the connection between features in a neural network to predict the human pose. Current neural network architectures that can be combined with the human skeleton approach are recurrent neural networks (RNN) [28, 29], long short-term memory networks (LSTM) [30, 31], convolutional neural networks (CNN) [32], etc. To make the human skeleton approach more general, [25, 26] proposed

to use the heat map as a complement to the skeleton information and to use the human pose image in each video frame for the encoding process. The feature communication between bone joint points is shown in Figure 1.

Literature [33] proposed a method to construct a human action dataset combining skeleton information with video in order to improve the pose estimation and action recognition accuracy of CNN networks. Literature [34] proposed a multitask parallel learning framework to improve the accuracy and stability of body joint detection. Literature [35] proposed a human intention algorithm aiming at learning behavioral action features through environmental assistance. Literature [36] took the approach of attention mechanism, which divides the human body into different parts and obtains attention from each part separately to recognize actions. Some researchers have found that the spatial-temporal graph convolution network (ST-GCN) can utilize the spatial-temporal information of skeletal articulation points effectively. It performs spatial-temporal convolution on the skeletal graph, models the graph representation of each skeleton, and uses a subsequent temporal filter to capture dynamic temporal information, as shown in Figure 2.

3. Method

3.1. Graph Convolutional Network. Benefiting from [37], the sequence of each frame t of the human skeleton in space is expressed as follows:

$$f_{out} = \sum_{d=0}^D W_d f_{in} (\Lambda_d^{-1/2} A_d \Lambda_d^{-1/2} \otimes M_d), \quad (1)$$

where D represents the maximum distance of the graph, f_{in} and f_{out} represent the input and output values of the feature map, \otimes represents the multiplication function, A and d mark the d -order adjacency matrix of the joint pair, and the result of the normalization operation is represented by A_d . W_d and M_d indicate adaptive adjustment parameters. It plays an important role in the realization of boundary adjustment and convolution operations. In order to extract temporal features, we insert a $L \times 1$ convolutional layer in the shallow layer to fuse the space information of the joint points between adjacent frames. In the process of temporal feature extraction, L represents the length of the time window, which is a predefined hyperparameter. Each time unit and space unit are followed by a BatchNorm module and a ReLU module to form a whole with this structure.

3.2. Residual Unit. Literature [38] proposed a structure called bottleneck, which cleverly uses the advantages of $\text{conv}1 \times 1$ and is placed in the front and back positions of the common convolution part to reduce the number of feature channels in the convolution operation. In this paper, we cleverly used the bottleneck structure, abandoning the original time and space modules, and found in the experiment that the improved structure is significantly faster in model training and parameter calculation. For example, the input and output channels are 256, the channel reduction rate $r=4$, and the time window size $L=9$. Then, the total

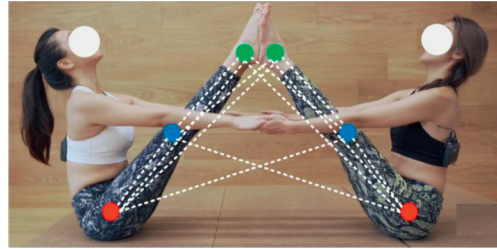


FIGURE 1: The feature communication between bone joint points.

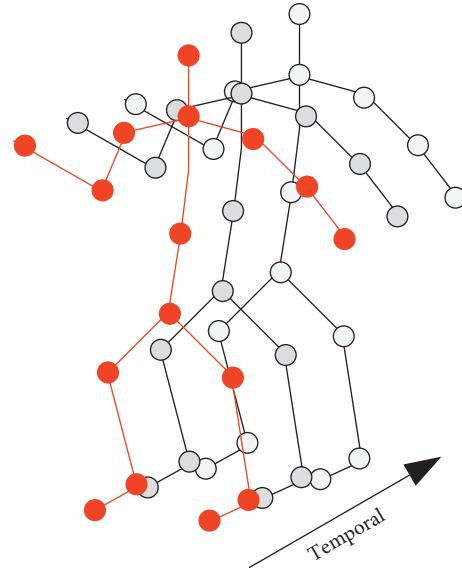


FIGURE 2: The principle of skeleton spatiotemporal feature extraction.

number of parameters involved in the calculation of the original structure is $256 \times 256 \times 9 = 589824$. If the bottleneck structure is adopted, the total number of parameters involved in the calculation is $256 \times 64 + 64 \times 64 \times 9 + 64 \times 256 = 69632$. Comparing the two, it can be seen that the bottleneck structure reduces the number of parameters calculated by the original structure by 8.5 times. Finally, we propose a new PartAtt block to enhance the generalization ability of the model. An example of a bottleneck structure frame is shown in Figure 3.

Considering that the time module and the space module in the original structure cannot integrate the features well, we connect the time and space modules with the residual structure to construct the ResGCN unit. The specific residual connection structure is shown in Figure 4. The Module residual module adopts a jump connection mode, the Block residual module adopts the mode of connecting before and after, the Dense residual module integrates the connection mode of the Module residual module and the Block residual module, making the structure more compact and saving calculation costs.

3.3. Multistream Input Structure. As we know from the bottleneck structure framework, each layer of input can be represented by a set of hyperparameters. In the first layer, we

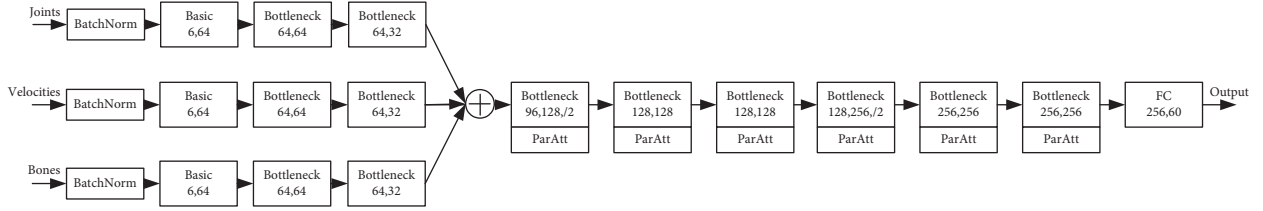


FIGURE 3: Yoga action recognition network fused with bottleneck structure.

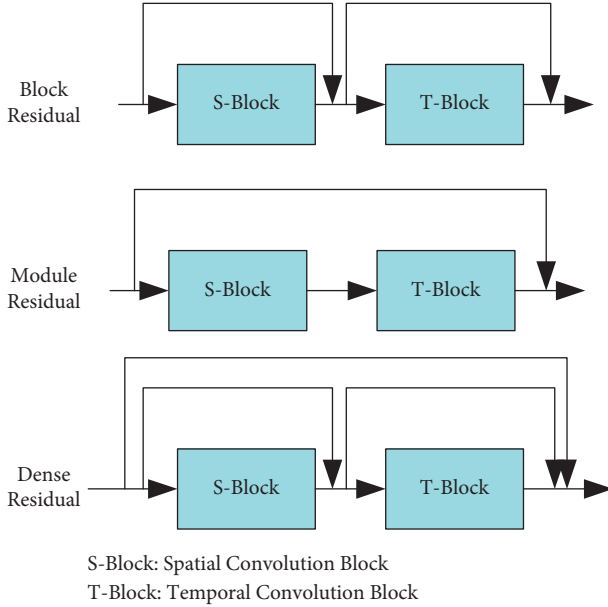


FIGURE 4: Three types of residual structure.

usually use basic operations to process the original input data. The second layer starts to design the bottleneck structure to filter the output data of the previous layer, and the difference in the design of the bottleneck structure is the different number of channels between the input and output. The third and fourth layers also use the bottleneck structure, but the only difference is that each layer is followed by a PartAtt unit. By introducing the PartAtt unit, all the position information of the extracted feature vector is preserved. In the decoding process, the encoding can be performed directly by the PartAtt mechanism, which reduces the intermediate steps of traditional decoding and solves the problem of feature loss. Secondly, in the PartAtt mechanism, each step of encoding and decoding directly accesses the source feature library, which realizes the direct feature tradition of encoding and decoding and shortens the exchange in feature transfer. In addition, the time step is set to 2 in the input stage of the third and fourth layers to further reduce the complexity of parameter computation and prevent over-fitting problems.

Furthermore, in high-precision models, input data generally require a multistream architecture for presentation. For example, the dual-stream input architecture mentioned in [39] incorporates both joint data and skeletal data as inputs, and decision selection is made after

multiple streams of inputs. This approach is adopted by most researchers because it is effective in improving model performance. However, the multistream architecture does not control the computational cost well, and the large amount is data input, parameter exchange, and variable calculation in the multistream framework, which invariably increases the huge computational volume. Therefore, our action recognition model adopts a multistream architecture in the pretraining stage, with a total of three input branches, and each input branch feature is fused with mainstream features in a pass-through tandem manner. This structure not only preserves the skeleton features to a great extent, but also makes the model more concise in its vertical structure and easier to converge when the model is trained.

In the data preprocessing stage, we mainly used the methods proposed in [29, 40] for reference. In the motion recognition method based on bone joint points, data preprocessing is very critical. In our work, preprocessing mainly revolves around joint positions, motion speeds, and bone characteristics. Suppose that a video of the action sequence is collected. According to the action sequence, the spatial coordinate set is $X = \{x \in \mathbb{R}^{C \times T \times V}\}$, where C represents the coordinates, T represents the frame, and V represents the joints. You can also get the set of relative positions of bones in space $R = \{r_i | i = 1, 2, \dots, V\}$, where $r_i = x[:, :, i] - x[:, :, c]$, $x[:, :, c]$ represents human bones and spinal joints. Combining the sets R and X into one set can be input into the multistream branching framework as the joint positions in action recognition. In addition, two sets of speeds of each joint can be obtained $F = \{f_t | t = 1, 2, \dots, T\}$ and $S = \{S_t | t = 1, 2, \dots, T\}$, where $f_t = x[:, t+2, :] - x[:, t, :]$ and $S_t = x[:, t+1, :] - x[:, :]$. Each motion feature of each joint can be represented by the two sets of feature vectors F and S , and this is input into the multistream branch frame as a motion stream. The basic characteristics of bones include length $L = \{L_i | i = 1, 2, \dots, V\}$ and angle $A = \{A_i | i = 1, 2, \dots, V\}$. The angle and length of the bone can be calculated through the bone displacement relationship $l_i = [x[:, :, i] - x[:, :, i_{adj}]]$, where the first joint of i_{adj} represents the adjacent joint. The calculation equation for the angle obtained by conversion of the customs clearance equation is as follows:

$$a_{i,w} = \arccos\left(\frac{l_i \cdot w}{\sqrt{l_{i,x}^2 + l_{i,y}^2 + l_{i,z}^2}}\right), \quad (2)$$

where $w \in \{x, y, z\}$ represents space coordinates.

3.4. Partial Perception Framework. Long short-term memory neural network (LSTM) was proposed by Hochreiter [41] in 1997.

LSTM is a derivative of Recurrent Neural Network (RNN). Since 2010, it has been proven that RNN has been successfully applied to speech recognition [42], language modeling [43], and text generation [44]. However, the disappearance of gradients and explosions makes RNN difficult to apply to long-term dynamics research. As an improved network of RNN, LSTM can handle this problem well. LSTM gives the network a lot of freedom, so that the network memory unit has an adaptive solution to learn and update information, which greatly improves the performance of some perception networks.

Assume that $X = (x_1, x_2, \dots, x_n)$ represents an input sentence composed of word representations of n words. In every position t , the RNN produces a hidden layer h in the middle denoted as y_t , and the hidden state h_t uses a non-linear activation function to update the previously hidden layers h_{t-1} and the input x_t , as shown below:

$$\begin{aligned} y_t &= \sigma(W_y h_t + b_y), \\ h_t &= f(h_{t-1}, x_t), \end{aligned} \quad (3)$$

where W_y and b_y are the parameter matrices and vectors learned during the training process, and σ represents the elementwise softmax function.

The LSTM unit includes an input gate i_t , a forget gate f_t , an output gate o_t , and a memory unit c_t to update the hidden state h_t , as shown below:

$$\begin{aligned} i_t &= \sigma(W_i x_t + V_i h_{t-1} + b_i), \\ f_t &= \sigma(W_f x_t + V_f h_{t-1} + b_f), \\ o_t &= \sigma(W_o x_t + V_o h_{t-1} + b_o), \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tanh(W_c x_t + V_c h_{t-1} + b_c), \\ h_t &= o_t \odot \tanh(c_t), \end{aligned} \quad (4)$$

where \odot is a kind of function which is similar to the multiplication operation, V represents a matrix related to weight, and b represents the learning vector. To increase the model's performance, morpheme training was carried out on two LSTMs. The first one is a morpheme that begins on the left and works its way to the right; the next one is a reverse duplicate of a character. Before passing to the next layer, the outputs of the forward and reverse passes are combined in series. Finally, the prediction value is observed using the activation function.

After understanding the partial perception algorithm LSTM, it was inspiring, because in the human body recognition process, the human skeleton will be divided into multiple parts. Each part is an interconnected joint. These parts composed of joints are made by hand, for the graph convolution to be able to explore the relationship between these parts and extract the corresponding spatial features of the joint points. To obtain the information of a point in GCN, it is necessary to start from the field of that point. According to the adjacency matrix in the field, the skeleton

data is automatically segmented, and then all the feedback information is input to the next joint point to complete the capture of the feature points of the entire human skeleton. Through this operation, the defects of manual design features are avoided, and the spatial features on the time series are obtained. [45].

If an ordinary convolutional neural network is used, all parts will be merged into a whole for feature extraction of convolution operations. Partial perception networks can divide joints into different departments and capture individual features for each part. Separately extracting features in this way helps to explore the connection between parts, that is, the spatial-temporal relationship between joints. The structure of our proposed spatial-temporal graph convolutional network-based yoga action recognition is shown in Figure 5.

4. Experiments

4.1. Data Collection. Before the establishment of the yoga posture database, we referred to yoga courses and training materials to find a reasonable grading system to assess the risk of yoga postures. As mentioned in [46, 47], the researcher compared the physical extensibility and commonality of action between the different postures. It was also approached in terms of breathing rate, posture intensity, and meditation. Also, we interviewed a yoga instructor who showed us all the standard yoga poses and broke down each pose. From his experience's we learned that currently there are 6 main yoga poses such as standing, forward bending, sitting, twisting, back bending, and supine. Each movement determines a different level of body stretch. In the study of this paper, the grading mainly revolves around these movements; our experimental scoring is based on the depth camera directly in front as the main interface. The specific grading is shown in Table 1.

In preparing the yoga dataset, we invited a professional instructor for standard yoga posture data collection. Then we invited participants who had one year of yoga experience and those who had no previous yoga experience to divide into two groups and complete each group of movements under the guidance of the instructor. In the process of data collection, not only the body posture but also the duration and number of movements of each yoga posture were collected. The yoga duration refers to the total time from when the breathing is adjusted until after the posture is completely relaxed. The number of movements refers to the sum of all the postures done during the training period, except for some correction of the postures by the instructor. The Azure Kinect DK was used to collect yoga movement data. The data is then manually calibrated by us after the data collection is completed and the data is split. In order to enhance the validity of the data, we added confidence parameters in the coding process. The data collection results are shown in Table 2.

4.2. Model Training. In the model training process, we used Pytorch to implement yoga movement recognition and

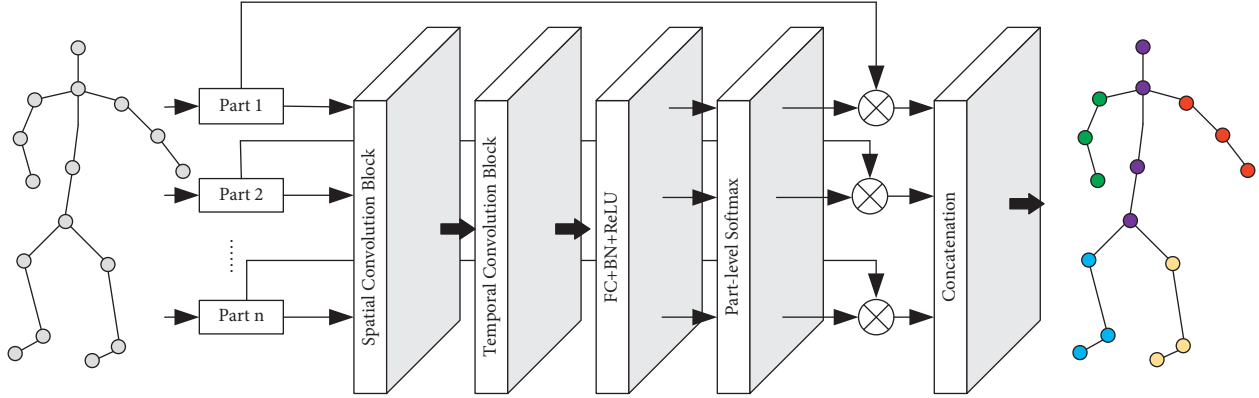


FIGURE 5: The overall network structure of yoga action recognition.

TABLE 1: Yoga action grading details.

Posture	Grade	Frontal	Sagittal
Standing	1	8	5
Sitting	2	6	6
Supine	3	10	9
Twisting	4	6	8
Forward bending	5	7	10
Back bending	6	9	6

TABLE 2: Yoga action time and frequency.

Posture	Experienced yogis		Inexperienced yogis	
	Ave time (s)	Ave frequency	Ave time (s)	Ave frequency
Standing	15	5	10	5
Sitting	31	6	16	6
Supine	23	8	14	8
Twisting	16	6	9	6
Forward bending	32	4	11	4
Back bending	27	5	12	5

grading. First, we used Openpose to extract the skeleton information from the yoga video dataset, and in each frame of the video we obtained the spatial coordinate information of each of the 14 joints. Then we use the heat map as the basis for pose estimation and perform secondary feature capture on the human skeleton. Then each frame of data is arranged in the temporal dimension to correlate the features between the joints from the temporal dimension. Finally, the skeletal joint features are fused using the average prediction score and the weights are estimated in a progressive ranking. We set different learning rates at different epochs. At the beginning of training, the learning rate is set to 0.05 to adapt to the training speed of the data. Then the learning rate is set to 0.01 at epoch = 30 to speed up the learning speed; after that, the learning rate is gradually reduced at epoch = 50 and epoch = 60 to find the optimal solution. The specific parameters in the model training are shown in Table 3. All the work is done in Ubuntu 16.04 and the whole training and prediction process is done with NVIDIA TITAN X GPU support on Intel Xeon E5-2620 CPU.

TABLE 3: Training parameter settings.

Parameter	Value
Epoch	20
Dropout rate	0.5
Initial learning rate	0.05
Learning rate (epoch = 30)	0.01
Learning rate (epoch = 50)	0.002
Learning rate (epoch = 60)	0.0004
Weight attenuation coefficient	0.0002
Momentum	0.9

4.3. *Experimental Result.* For the experimental data collection, we collected 50 experienced yogis and 50 inexperienced yogis. And the data was split according to the previous solution. The sensitivity, specificity, precision, and accuracy of skeletal features were captured in the data in the split starting from each frame. The experimental results are shown in Table 4. The standard yoga movements were decomposed on a larger scale, making it traceable in the validation set. Based on the above statistical results, higher

TABLE 4: Yoga action recognition results.

Posture		Sensitivity	Specificity	Precision	Accuracy
Standing	Experienced yogis	0.98	0.99	0.95	0.99
	Inexperienced yogis	0.71	0.91	0.91	0.98
Sitting	Experienced yogis	0.89	0.98	0.93	0.98
	Inexperienced yogis	0.66	0.91	0.91	0.98
Supine	Experienced yogis	0.94	0.99	0.89	0.99
	Inexperienced yogis	0.78	0.93	0.91	0.97
Twisting	Experienced yogis	0.96	0.98	0.91	0.99
	Inexperienced yogis	0.77	0.94	0.89	0.99
Forward bending	Experienced yogis	0.93	0.99	0.87	0.99
	Inexperienced yogis	0.69	0.91	0.89	0.98
Back bending	Experienced yogis	0.91	0.99	0.92	0.97
	Inexperienced yogis	0.72	0.92	0.86	0.97

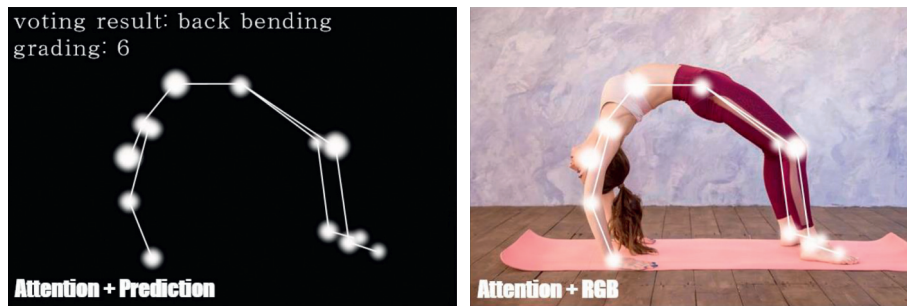


FIGURE 6: The effect of yoga action recognition and grading system.

TABLE 5: Yoga action grading results.

Posture	Grade	Experienced yogis (%)	Inexperienced yogis (%)
Standing	1	91	80
Sitting	2	93	77
Supine	3	89	75
Twisting	4	93	76
Forward bending	5	93	71
Back bending	6	92	72
Total ave	90	75	

sensitivity values represent more experience in yoga training and also predict closer standardization of yoga poses.

From the above experimental results, we can see that the recognition accuracy of all yoga poses is close to 1. And the accuracies, as a kind of random error, all keep above 0.86, which proves that the model performance is still great. The gap between experienced yogis and inexperienced yogis is mainly in sensitivity and specificity. Experienced yogis scored higher in both metrics, representing the more standardized yoga poses. The yogis are captured by the depth camera while practicing yoga. Real-time skeletal joint tracking is performed on the captured video. Finally, the yoga movements are recognized with the training model and then matched with the database to generate a grading score. The specific recognition effect is shown in Figure 6.

In addition, we also made corresponding statistics in the grading, as shown in Table 5.

Table 5 demonstrates that the average grading accuracy of experienced yogis in the whole set of yoga poses is higher than that of inexperienced yogis. The yoga posture with the

greatest difference was forward bending, followed by back bending. Because of the difficulty of these two poses, it was difficult for inexperienced yogis to achieve the standard poses, so the accuracy of poses grading was lower. The above experimental results favorably prove the effectiveness of the grading system in this paper, which can give yogis feedback and remind them to change their postures if the yoga movements are not standard.

5. Conclusion

In this paper, we found that, with the popularity of the Internet, people's lifestyles have also changed, and many people choose to learn yoga by watching videos on the Internet. For yoga beginners, learning yoga online in this way without the direct guidance of an instructor, there is a high chance that the yoga poses will be substandard. Highly difficult yoga poses are likely to be disabling for beginners. To address this potential risk, we propose a yoga posture recognition and grading system based on spatial-temporal

graph convolutional neural network. We first use LSTM network to label yoga practice videos frame by frame. Then we extract the skeletal joint point features sequentially with graph convolution and then obtain the connection between joints from arranging each video frame in spatial-temporal dimension and correlating the joint points in each frame with neighboring frames for spatial-temporal information. Finally, through experiments, it is proved that our method can accurately identify yoga poses and grade them accordingly and can identify whether the yoga poses are standard or not and at the same time give feedback to yogis in a timely manner to prevent injuries to the body caused by nonstandard poses.

For deep learning algorithms, the larger the number of datasets, the better the accuracy of the model obtained from training. Since there is no specific dataset for yoga poses at present, the number of homemade datasets in this paper is small, which is the shortcoming of the work in this paper. Making datasets is a tedious and time-consuming task. In our future work, we will gradually increase the number of datasets, and at the same time, we will invest more efforts in the field of data preprocessing.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper was supported by science and technology research project of Colleges and Universities in Inner Mongolia Autonomous Region, Project no. NJSY22477.

References

- [1] Yogi Times, "Demographics & statistics of the yoga industry December 28," 2020, <https://www.yogitimes.com/article/unstoppable-trend-yoga-infographic-business>.
- [2] P. Salmon, E. Lush, M. Jablonski, and S. E. Sephton, "Yoga and mindfulness: clinical aspects of an ancient mind/body practice," *Cognitive and Behavioral Practice*, vol. 16, no. 1, pp. 59–72, 2009.
- [3] J. B. Webb, C. B. Rogers, and E. V. Thomas, "Realizing Yoga's all-access pass: a social justice critique of westernized yoga and inclusive embodiment," *Eating Disorders*, vol. 28, no. 4, pp. 349–375, 2020.
- [4] Roy Morgan Research, "Yoga participants profile," 2016, <https://www.roymorgan.com/findings/7004-yoga-is-the-fastest-growing-sport-or-fitness-activity-in-australia-june-2016-201610131055>.
- [5] Roy Morgan Research, "Yoga participation stretches beyond Pilates and aerobics March 30," 2018, <https://www.roymorgan.com/findings/7544-yoga-pilates-participation-december-2017-201803290641>.
- [6] Ipsos Public Affairs, "2016 yoga in America study," 2017, <https://www.yogaalliance.org>.
- [7] J. K. Thompson, L. J. Heinberg, M. Altabe, and S. Tantleff-Dunn, "The scope of body image disturbance: the big picture [J]," *Exacting beauty: Theory, Assessment, and Treatment of Body Image Disturbance*, American Psychological Association, United States, pp. 19–50, 1999.
- [8] A. Borden and C. Cook-Cottone, "Yoga and eating disorder prevention and treatment: a comprehensive review and meta-analysis," *Eating Disorders*, vol. 28, no. 4, pp. 400–437, 2020.
- [9] E. A. Impett, J. J. Daubenmier, and A. L. Hirschman, "Minding the body: yoga, embodiment, and well-being," *Sexuality Research and Social Policy*, vol. 3, no. 4, pp. 39–48, 2006.
- [10] L. Mahlo and M. Tiggemann, "Yoga and positive body image: a test of the Embodiment Model," *Body Image*, vol. 18, pp. 135–142, 2016.
- [11] N. Piran and D. Neumark-Sztainer, "Yoga and the experience of embodiment: a discussion of possible links," *Eating Disorders*, vol. 28, no. 4, pp. 330–348, 2020.
- [12] A. E. Cox and T. L. Tylka, "A conceptual model describing mechanisms for how yoga practice may support positive embodiment," *Eating Disorders*, vol. 28, no. 4, pp. 376–399, 2020.
- [13] M. Singleton, *Yoga Body: The Origins of Modern Posture Practice*, Oxford University Press, United Kingdom, 2010.
- [14] A. R. Jain, *Selling Yoga: From Counterculture to Pop Culture*, Oxford University Press, United Kingdom, 2015.
- [15] J. B. Webb, E. R. Vinoski, J. Warren-Findlow, M. I. Burrell, and D. Y. Putz, "Downward dog becomes fit body, inc.: a content analysis of 40 years of female cover images of Yoga Journal," *Body Image*, vol. 22, pp. 129–135, 2017.
- [16] N. Bhalla and D. Moscovitz, "Yoga and female objectification: commodity and exclusionary identity in U.S. Women's magazines," *Journal of Communication Inquiry*, vol. 44, no. 1, pp. 90–108, 2020.
- [17] C. Sminchisescu, A. Kanaujia, Z. Li, and D. Metaxas, "Discriminative density propagation for 3d human motion estimation[C]," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 390–397, IEEE, San Diego, CA, USA, June 2005.
- [18] R. Rosales and S. Sclaroff, "Learning body pose via specialized maps[C]," in *Proceedings of the Advances in Neural Information Processing Systems 14*, pp. 1263–1270, Vancouver, British Columbia, Canada, December 3–8, 2001.
- [19] A. Agarwal and B. Triggs, "Monocular human motion capture with a mixture of regressors[C]," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, p. 72, September 2005.
- [20] L. Ren, G. Shakhnarovich, J. K. Hodgins, H. Pfister, and P. Viola, "Learning silhouette features for control of human motion," *ACM Transactions on Graphics*, vol. 24, no. 4, pp. 1303–1331, 2005.
- [21] I. Cohen and H. Li, "Inference of human postures by classification of 3D human body shape[C]," in *Proceedings of the 2003 IEEE International SOI Conference. Proceedings (Cat. No. 03CH37443)*, pp. 74–81, IEEE, Nice, France, October 2003.
- [22] R. Rosales, M. Siddiqui, J. Alon, and S. Sclaroff, "Estimating 3D body pose using uncalibrated cameras[C]," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, p. I, December 2001.
- [23] K. Grauman, G. Shakhnarovich, and T. Darrell, "Inferring 3D structure with a statistical image-based shape model[C]," in

- Proceedings of the Ninth IEEE International Conference on Computer Vision*, p. 641, Nice, France, October 2003.
- [24] X. Chen and M. Koskela, "Skeleton-based action recognition with extreme learning machines," *Neurocomputing*, vol. 149, pp. 387–396, 2015.
- [25] M. Liu and J. Yuan, "Recognizing human actions as the evolution of pose estimation maps[C]," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1159–1168, Salt Lake City, UT, USA, June 2018.
- [26] W. Du, Y. Wang, and Y. Qiao, "Rpan: an end-to-end recurrent pose-attention network for action recognition in videos[C]," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3725–3734, Venice, Italy, October 2017.
- [27] J. Yu, M. Jeon, and W. Pedrycz, "Weighted feature trajectories and concatenated bag-of-features for action recognition," *Neurocomputing*, vol. 131, pp. 200–207, 2014.
- [28] S. Song, C. Lan, J. Xing, W. Zeng, and J. Liu, "An end-to-end spatio-temporal attention model for human action recognition from skeleton data[C]," *Proceedings of the AAAI conference on artificial intelligence*, vol. 31, no. 1, pp. 4263–4270, 2017.
- [29] C. Si, Y. Jing, W. Wang, L. Wang, and T. Tan, "Skeleton-based action recognition with spatial reasoning and temporal stack learning[C]," in *Proceedings of the European Conference on Computer Vision (ECCV)*, Springer, Cham, pp. 103–118, 2018.
- [30] W. Zhu, C. Lan, J. Xing, and Y. Li, "Co-occurrence feature learning for skeleton based action recognition using regularized deep LSTM networks[C]," *Proceedings of the AAAI conference on artificial intelligence*, vol. 30, no. 1, 2016.
- [31] J. Liu, A. Shahroudy, D. Xu, and G. Wang, "Spatio-temporal LSTM with trust gates for 3D human action recognition," in *Proceedings of the European conference on computer vision*, Springer, Cham, pp. 816–833, 2016.
- [32] C. Li, Q. Zhong, D. Xie, and S. Pu, "Co-occurrence feature learning from skeleton data for action recognition and detection with hierarchical aggregation[J]," arXiv preprint arXiv:1804.06055, 2018.
- [33] U. Iqbal, M. Garbade, and J. Gall, "Pose for action-action for pose[C]," in *Proceedings of the 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pp. 438–445, IEEE, Washington, DC, USA, June 2017.
- [34] D. C. Luvizon, D. Picard, and H. Tabia, "2d/3d pose estimation and action recognition using multitask deep learning [C]," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5137–5146, Salt Lake City, UT, USA, June 2018.
- [35] B. Xu, J. Li, Y. Wong, Q. Zhao, and S. M. Kankanhalli, "Interact as you intend: intention-driven human-object interaction detection[J]," *IEEE Transactions on Multimedia*, vol. 22, no. 6, pp. 1423–1432, 2019.
- [36] B. Wan, D. Zhou, Y. Liu, R. Li, and X. He, "Pose-aware multi-level feature network for human object interaction detection [C]," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9469–9478, Seoul, Korea (South), November 2019.
- [37] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition [C]," in *Proceedings of the Thirty-second AAAI conference on artificial intelligence*, Louisiana, New Orleans, USA, February 2018.
- [38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition[C]," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [39] L. Shi, Y. Zhang, J. Cheng, and H. Lu, "Two-stream adaptive graph convolutional networks for skeleton-based action recognition[C]," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12026–12035, Long Beach, CA, USA, June 2019.
- [40] Y. F. Song, Z. Zhang, and L. Wang, "Richly activated graph convolutional network for action recognition with incomplete skeletons[C]," in *Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP)*, pp. 1–5, IEEE, Taipei, Taiwan, September 2019.
- [41] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [42] O. Vinyals, S. V. Ravuri, and D. Povey, "Revisiting recurrent neural networks for robust ASR[C]," in *Proceedings of the 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 4085–4088, IEEE, Kyoto, Japan, March 2012.
- [43] T. Mikolov, M. Karafiát, L. Burget, J. Černocký, and S. Khudanpur, "Recurrent neural network based language model[C]," *Interspeech*, vol. 2, no. 3, pp. 1045–1048, 2010.
- [44] I. Sutskever, J. Martens, and G. E. Hinton, "Generating text with recurrent neural networks[C]," in *Proceedings of the 28th International Conference on Machine Learning*, Bellevue, Washington, USA, June 2011.
- [45] J. Liang and G. Zuo, "Taekwondo action recognition method based on partial perception structure graph convolution framework[J]," *Scientific Programming*, vol. 2022, pp. 1–10, 2022.
- [46] U. H. Graneheim and B. Lundman, "Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness," *Nurse Education Today*, vol. 24, no. 2, pp. 105–112, 2004.
- [47] M. Bengtsson, "How to plan and perform a qualitative study using content analysis," *NursingPlus Open*, vol. 2, pp. 8–14, 2016.

Research Article

China's Economic Forecast Based on Machine Learning and Quantitative Easing

Chang Qiu 

School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China

Correspondence should be addressed to Chang Qiu; 2015311044@live.sufe.edu.cn

Received 11 February 2022; Accepted 3 March 2022; Published 26 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Chang Qiu. 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, six variables, including export value, real exchange rate, Chinese GDP, and US IPI, and their seasonal variables, are used as determinants to model and forecast China's export value to the US using three methods: BP neural network, ARIMA, and AR-GARCH. Error indicators were chosen to compare the simulated and predicted results of the three models with the real values. It is found that the results of all three models are satisfactory, although there are some differences in their simulation and forecasting capabilities, but the ARIMA model has a clear advantage. This paper analyses the reasons for these results and proposes suggestions for improving China's exports in the context of the models.

1. Introduction

Export trade is one of the driving forces behind China's rapid economic growth, and the United States, as China's top export trading partner, has a huge impact on China's economic development. However, recently, due to the global economic crisis and the appreciation of the RMB against the USD, China's export growth has started to slow down and even become negative. Therefore, modelling China's exports to the US and quantifying the factors that influence them is extremely important in order to predict and take measures to increase China's export value [1].

Quantitative easing has been in place for many years, but the theoretical basis for this so-called unconventional policy has not yet been agreed upon by academics, and there is still much disagreement. In addition to the "liquidity trap theory," many believe that the "non-neutral money theory," "financial accelerator theory," and the "great depression theory" proposed by Bernanke and others are the theoretical foundations of this policy. The "great depression theory" proposed by Bernanke and others also forms the theoretical basis for quantitative easing [2]. As a result of the implementation of quantitative easing, his theories are considered to be the veritable foundation of quantitative easing. However, the relationship between MMT and quantitative

easing has been highlighted by the academic hype about modern monetary theory (MMT), the great recession and the implementation of unlimited quantitative easing, i.e., the former is the theoretical basis of the latter and the latter is the concrete practice of the former [3]. At present, academics have basically reached a consensus on whether MMT belongs to the theoretical basis of quantitative easing policy, but there is a greater controversy over the understanding of quantitative easing policy itself and its theoretical basis. Some domestic scholars who hold opposing views believe that MMT has major theoretical flaws and its basic proposition is a serious departure from reality, and thus a fallacy. The quantitative easing policy based on this theory also has major flaws, as it is essentially a kind of monetisation of fiscal deficits, in which fiscal expenditure is regarded as money creation and fiscal revenue is regarded as money recovery [4]. This is essentially an act of "robbing the poor to give to the rich" and "pulling the wool over the world's eyes." Some mainstream economists abroad have also taken a negative view of MMT, arguing that it is both detrimental to the development of economic theory and harmful to real economic policy. Some scholars, however, are enthusiastic about the theoretical claims of MMT, arguing that it is a disruptive theoretical innovation, a fundamental rejection of mainstream

macrofinance and fiscal theory, and even a solution to a major problem in the field of political economy.

The author believes that whether the theoretical basis of quantitative easing policy is scientific or not must not only be tested by practice but also evaluated from the perspective of Marxist political economy. Therefore, based on the effectiveness of quantitative easing policy in the US, this paper systematically analyses the relationship between the operation process of quantitative easing policy and its theoretical basis, discusses the trend of normalisation of quantitative easing policy and its limitations from the perspective of Marxist political economy, makes necessary arguments on the internal link between quantitative easing policy and MMT, and points out the importance of conducting this study to China's understanding of quantitative easing policy and the formulation of policies to cope with the current round of the great recession. The study is also useful for China to understand quantitative easing and formulate policies to cope with the current round of the great recession [5].

2. Related Work

In the existing export trade forecasting literature, the main quantitative forecasting methods are ARIMA models, AR-GARCH models, neural network modelling methods, and some extensions based on them. Of these, the ARIMA and AR-GARCH models are upgrades of traditional time series models (exponential smoothing, moving average, etc.), and these linear forecasting methods have the advantage of being simple, intuitive, and highly explanatory. There is also a small body of literature on export trade forecasting using cointegration models, support vector machine models, and grey system methods. In terms of variable selection, the usual empirical variables are often used: export value variables, real exchange rate variables, GDP variables of importing and exporting countries, and their lagged variables. In addition, since macrovariables tend to be highly seasonally cyclical, they are generally seasonally adjusted in the literature [6].

The empirical evidence is as follows: Yildirim and Ivrendi [7] extend the univariate ARMA model to multiple variables and test the model's effectiveness with forecasts of export effects for Sweden. A univariate ARIMA model based on a Bayesian search algorithm is presented in [7], which is empirically shown to be less effective in predicting seasonal models than nonseasonal models. In [8], a neural network model was used to forecast the export of DOC in Scotland, and methods such as increasing the sample size and improving the parameters were proposed to improve the accuracy of the model.

The ARIMA model was developed using the monthly export data of China from 2015 to 2021, with lagged export data as the determinant variable, trend differencing and seasonal differencing, and was used to forecast China's exports in 2023. [9] An AR-GARCH model was developed to simulate China's exports to the US using empirical variables as determinants and excluding seasonal effects, and a dynamic conditional correlation coefficient (DCC) model was

used to analyse the impact of RMB appreciation on exports. Fofack et al. [10] analyse the application of cointegration techniques and error correction models based on them in forecasting, propose a nonlinear error correction forecasting model based on neuronal networks, combining the characteristics of neuronal networks, and forecasts the general situation of China's export trade in 2023. [11] An empirical analysis of Hunan Province's exports was conducted and the model was found to be effective in forecasting regional export trade.

In a comparative analysis of these methods, Cole et al. [12] used a neural network model, exponential smoothing, and ARIMA to forecast Thai rice exports and evaluated the effectiveness of the three methods using various indicators. Woo and Zhang [13] compared the advantages and disadvantages of the ARCH family model and the BP neural network model in predicting maize and stock prices, and both concluded that the nonlinear system modelling was slightly better than the linear model.

3. Basic Principles of the Three Models

3.1. Basic Principles of BP Neural Networks. A BP network is a multilayer feed-forward network with backpropagation of errors and is the most representative and widely used type of an artificial neural network. To train a BP network, the same set of inputs and desired outputs are used as training "samples," and the network is trained according to a certain algorithm, and once trained, the model can be used to solve similar problems. A BP network requires a training set and a test set to evaluate its training results. The former is used to train the network to achieve a specified error, and the latter is used to evaluate the performance of the trained network.

3.2. Fundamentals of the ARIMA Model. The ARIMA model, also known as the differential autoregressive moving average model, is an extension of the ARMA (p, q) model. In ARIMA (m, d, n), AR is the 'autoregressive,' m is the number of autoregressive terms; MA is the 'sliding average,' n is the number of sliding average terms, and d is the number of differences (orders) made to make it a smooth series. After d differences, the ARIMA (p, d, q) model can be expressed as an ARMA model with the following expression:

$$r_t = c + \sum_{i=1}^m \phi_i r_{t-i} + \sum_{j=1}^n a_j y_{t-j} + \varepsilon_t, \quad (1)$$

where r_{t-i} is the r_t lagged i-order variable, y_{t-j} is the moving average term at lag j, and ε_t is the residual, which follows a standard normal distribution.

3.3. Fundamentals of the AR-GARCH Model. To describe and predict volatility clustering in economic time series, an autoregressive conditional heteroskedasticity model (i.e. ARCH model) is used. Since the ARCH model is a short memory process, in order to better characterise certain financial market phenomena with long memory processes, the authors of [14] generalises the ARCH model and adds lags to

the residual term conditional variance to derive a generalised autoregressive conditional heteroskedasticity model (GARCH model). AR-GARCH is the addition of an autoregressive term to the mean expression of the GARCH model, which can be expressed as

$$\begin{aligned} r_t &= c + \sum_{i=1}^m \phi_i r_{t-i} + \varepsilon_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^p \alpha_i \sigma_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \end{aligned} \quad (2)$$

where r_t is the series of returns, σ_t^2 is the series of variances, and ε_t is the residuals, and it follows a standard normal distribution.

4. Application of the Model and Empirical Analysis

4.1. Selection of General Variables. The most commonly used explanatory variables for the export function in the literature on exchange rate exports are: real exports with a lag; the real exchange rate; the GDP of the exporting country, which is a measure of the exporting country's export capacity; and the GDP (gross domestic product) of the importing country or other proxy variables (such as the industrial production index IPI), which are a measure of the importing country's import capacity.

The lagged real exports are usually selected with a one-period lag, and in this paper, we also select a one-period lagged variable for China's exports to the US.

The real exchange rate is calculated using the formula

$$\text{real exchange rate} = \frac{\text{nominal exchange rate} \times \text{foreign price}}{\text{local currency commodity price}}. \quad (3)$$

Foreign prices and local currency commodity prices are expressed in terms of the consumer price index (CPI) for both countries. The GDP data for China and the US are actually only available on a quarterly basis, and the sample for this paper is based on a monthly interval (if the interval is annual, then the data can be collected over a shorter period of time and the sample size is smaller), so I have tried to replace the data with monthly IPI data. As there has been no IPI data for China since 2006, this paper can only average the quarterly GDP data for China over three months to obtain monthly GDP data, while the IPI data are used to replace the US GDP data.

4.2. Selection of Seasonal Variables. As macrodata tends to be highly seasonal, the graphical analysis of the sample series shows that there is strong seasonality in the value of exports to the US and China's GDP. Exports, for example (Figure 1), have a strong annual cycle of their own, with the value of exports generally increasing, but suddenly decreasing in January and February each year, and then increasing again. China's GDP is also seasonal, and this is more pronounced in China (Figure 1). This is strongly related to the Chinese

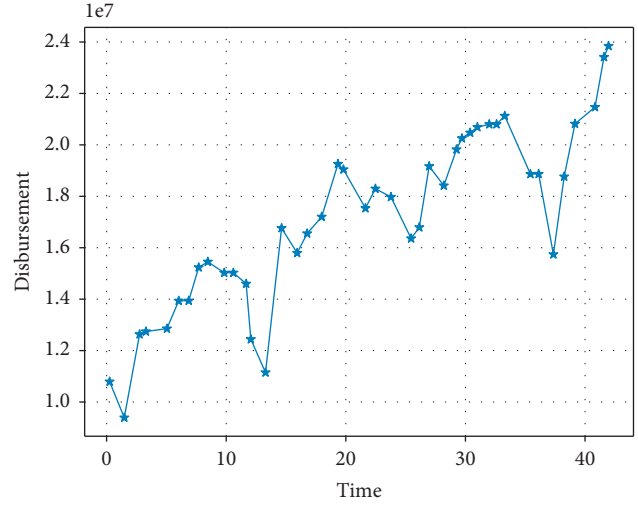


FIGURE 1: Data on China's exports to the US over the sample period (\$000s).

New Year holiday. So, in the forecasting model, we add seasonal lags to these two variables, i.e., the value of exports in the same period last year and the value of Chinese GDP in the same period last year, as determinants. This allows the model to capture the impact of seasonality on exports to the US and China's GDP.

In forecasting models, the dependent variable is usually replaced by a number of periods lagged, as data for the current independent variable is often not yet available for the forecast period. Therefore, the values of all decision variables in this model are brought into the model with a one-period lag. The six input variables of the input layer of this neural network prediction model are summarised in Figure 2.

4.3. Data Sources and Data Processing. Monthly data from 2015 to 2021 are used as sample data to model China's exports to the US. The nominal exchange rate of the RMB to the USD is obtained from the website of the State Administration of Foreign Exchange (SAFE); data on China's export trade to the US and China's Consumer Price Index (CPI) are obtained from the RESET database, while data on the US CPI and the Industrial Price Index (IPI) are obtained from the US Bureau of Labor Statistics and the Federal Reserve website.

To calculate the real exchange rate, the ratio of the CPI indices for China and the US over the sample period is required. Since the US CPI is a chain index and the Chinese CPI (used in this paper) is a year-on-year index, we first convert the CPI data for each country into a CPI-based index based on the CPI index of January 2005.

4.4. Development of a BP Neural Network Model for Export Forecasting. There are three main stages in building a neural network model and completing training and learning: the configuration stage, the training stage, and the output stage.

We have already identified the factors that influence export value, with six main variables to be considered

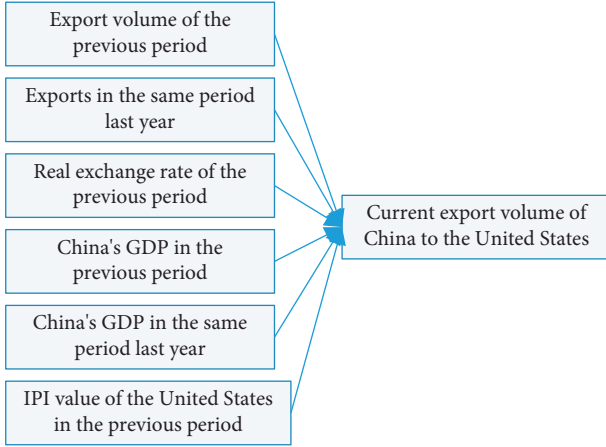


FIGURE 2: Summary of the variables determining the value of China's exports to the US.

(Figure 2), which we will use as input nodes to the BP model. The input nodes are normalised to the indicator data by the following formula:

$$X' = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})}. \quad (4)$$

The number of hidden layers and the number of nodes in the hidden layer are chosen as one layer. The number of nodes is directly related to the number of input and output units, which is determined by the formula $n1 = \sqrt{n + m} + a$, where m is the number of input neurons equal to 6, n is the number of output neurons of 1, and a is a constant between 1 and 10. In this paper, after repeated debugging, the number of nodes in the hidden layer was determined to be 10.

Output node is the variable to be predicted, the value of China's exports to the US, which is also normalised here.

This stage completes the training of the network on the sample. For the input information, the neural network is propagated forward to the nodes in the implicit layer, and then transferred to the output node after a sigmoid-type activation function. The sigmoid function expression is $y = 1/(1 + e^{-x/\beta})$ adaptive change according to the sample.

The rules were trained using the MATLAB `trainlm` function, and the Levenberg–Marquardt rules were used to train the forward network. The absolute mean percentage error (MAPE) was used as the error criterion for testing the samples. As mentioned earlier, the BP model was built with 6

input neurons, 10 hidden layer neurons, and 1 output neuron; the learning step was 0.06, the number of training sessions was 1,000, and the acceptable error criterion was $\varepsilon_0 = 0.001$.

5. Case Studies

However, since the model has a 12-period lag in the decision variable, the valid sample is only 32 periods. The data relating to exports within the sample period are used as the training sample and the test sample (the first 26 periods are used as the training sample, which satisfies the sample size requirement of $3 * k + 8$, with k being the number of explanatory variables of 6. The last 6 periods are used as test samples).

After inputting the samples, the system learns by minimising the sum of squared errors between the desired output and the actual output, adjusting the weight matrix and the threshold vector. After 20 training sessions, the error of the model is reduced to within the required range and the system stops learning. By running the MATLAB program 1,000 times, the best network (i.e. the one with the lowest error) for the detection sample is obtained and used as the final model for the neural network method. The prediction results for the detection samples are shown in Figures 3 and 4. Figure 5 shows that the model's predictions for China's exports to the USA are very close to the actual values, and the results are satisfactory. The average error of the normalized prediction sample is 0.0753, and the average error of the reduced prediction is 0.0365, which means that the theoretical deviation from the true value of exports predicted by the neural network method is no more than 3.65%.

Using this neural network, the forecast for September 2008 was US\$22,967,000 thousand, which is 6.821% less than the true value of US\$24,683,579 thousand. The forecast results are good.

6. Experimental Analysis

The ARIMA model is estimated based on the decision variables designed above. The first step is to check whether the original series of export values is stationary using the unit root test (ADF test). The results show that the export value series is smooth after first order differencing. Using the period January 2005 to August 2008 as the sample period, the model is estimated using EViews software as follows:

$$\begin{aligned} \Delta \text{EXPORT}_t = & 143796.03 - 0.38 * \Delta \text{EXPORT}_{t-1} (0.679) (0.020), \\ & \Delta ER_{t-1} (-1) + 39.90 * \Delta \text{GDP}_{t-1} (-1) + 29.34 * \Delta \text{GDP}_{t-1} (-12) - 395211.46 *, \\ & \Delta \text{IPI}_{t-1} (-1) + \varepsilon_{t-1} (0.011) (0.108). \end{aligned} \quad (5)$$

Here, ε_{t-1} is the error term, which follows the standard normal distribution. The p values of the coefficients are in parentheses.

The p values of the leading coefficients of $\Delta \text{EXPORT}_{t-1}$ and ΔGDP_{t-1} (0.050 and 0.108, respectively) are statistically

significant at the 90% confidence level, suggesting that they should be used as determinants of the model. Other variables with larger p values are not statistically significant, but empirical experience shows that their impact on export value is not negligible.

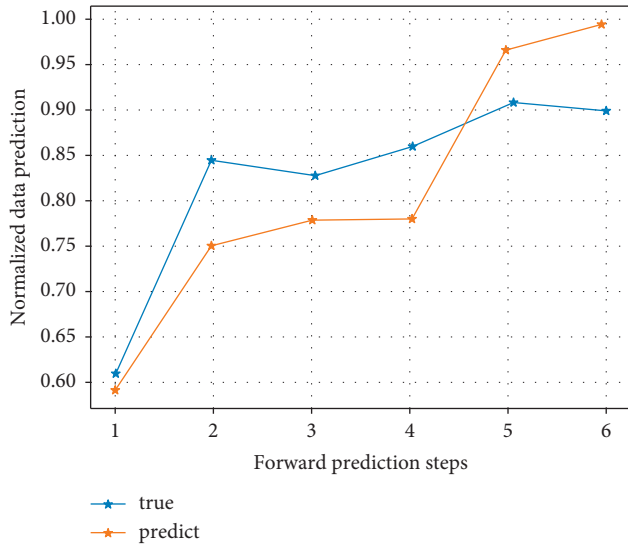


FIGURE 3: Export prediction.

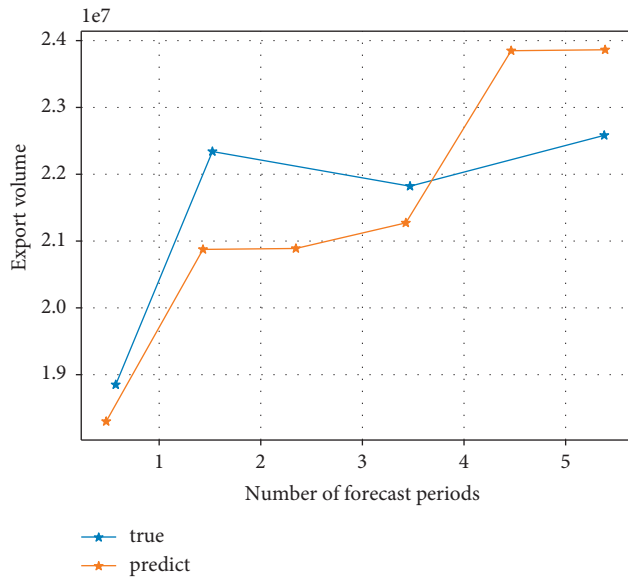


FIGURE 4: Comparison between the predicted and actual export values after reduction in the detection period of the neural network model export prediction versus actual value.

Using the above model to forecast the value of China’s exports to the USA over the sample period, the static forecast in EViews is used. Figure 6 shows a comparison between the predicted and true series of export values. It can be seen that the estimated ARMA model gives a better estimate of export values as the predicted values are derived from lagged decision variables. The mean absolute error of the model over the sample period is 5.503144%, which is greater than 3.65% predicted by the neural network approach for the test sample, which shows that the neural network approach outperforms the results predicted by the ARIMA model.

Based on this model, we can forecast one-period ahead, i.e., the export value for September 2008, using the sample data, and the forecast result is US\$ 248,248,864,824, while the real export value is US\$ 24,683,579,000, with an absolute

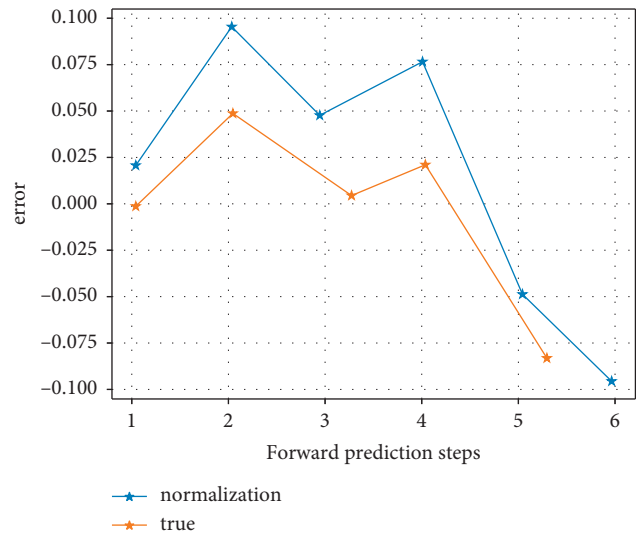


FIGURE 5: Normalized predicted values for the detection period of the neural network model and after reduction predicted versus actual values.

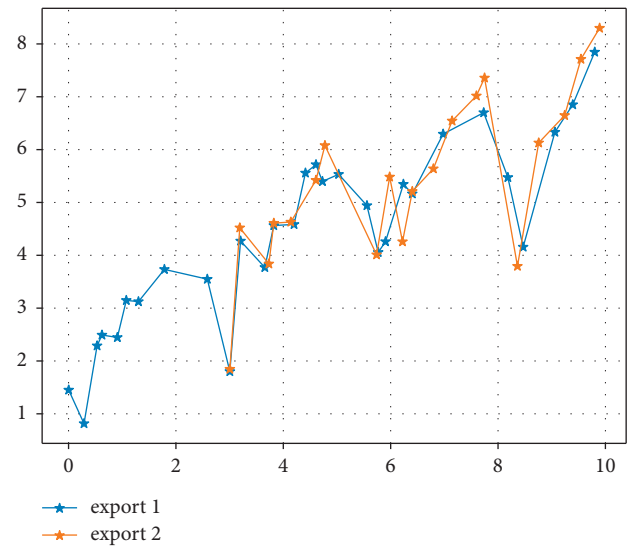


FIGURE 6: ARIMA model sample period exports.

error of 0.569%, which is very good. The ARIMA model outperforms the neural network approach in terms of this predictor.

Figure 7 shows a comparison of the AR(1)-GARCH model’s predicted and real true value series for exports over the sample period, which is broadly similar to the ARIMA model. Here, the results are not analysed in detail.

The results of the three models are compared. The absolute mean percentage error (MAPE) is used as a comparison criterion to evaluate the fit and prediction of the three models:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\text{Export forecast value} - \text{export actual value}}{\text{Actual export value}} \right| \quad (6)$$

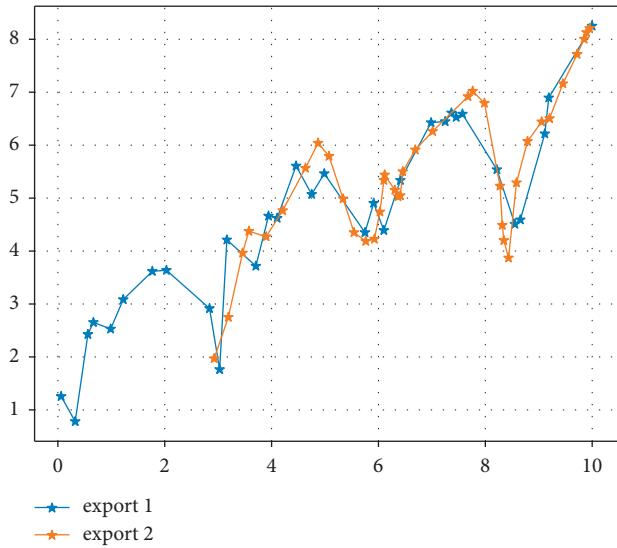


FIGURE 7: AR (1)-GARCH model sample period comparison of real and predicted export value series in the sample period.

The prediction results of the three methods were compared with the true values for the six periods selected in the sample period and are tabulated as shown in Table 1.

A comparison of the forecast errors for each period for the three methods is tabulated in Table 2.

These two tables allow a comparison of the results of the three models, with the following conclusions. The BP neural network model best fits the export value in the test period with an average error of 3.65%, while the ARIMA and AR-GARCH models have similar results with an average error of 5.5% over the sample period, which is worse than the neural network method. However, the latter two models perform very well in terms of predicting export value over the forecast period, with a prediction error of 0.614% compared to 6.821% for the neural network method [15]. This is contrary to the thesis that the neural network method outperforms the time series method in both rice export and stock price forecasting. The author believes that this is due to the fact that different papers have studied different subjects, adopted different variables, and built different models. In the other papers, the nonlinearity of the neural network method, which does not take into account the seasonal term, is found to explain the fluctuation of export value better than the general linear method. As this paper takes into account the seasonal variables more than in the previous literature, it proposes to add lagged seasonal variables to the model to deal with seasonal effects so that the linear model is able to describe and fit the intrinsic regularity of export movements very well, while the neural network approach produces results that are more random and contingent than the linear model. In addition, as mentioned above, the ARIMA model based on the Bayesian search algorithm is less effective in predicting seasonal models than nonseasonal models, and it is suggested that the seasonal treatment in this paper can be used to compare the forecasting effectiveness of the models.

TABLE 1: Comparison of predicted true values for each period obtained by the three methods.

	True value	Neural network method	ARIMA method	AR-GARCH method
2006.2	11218768	11948353	12060875	12135737
2007.2	16336347	17200569	15478242	16145578
2008.2	15475241	16253954	16754284	16886852
2008.5	21214314	21778439	21455741	27456721
2008.8	24035274	22788427	24451470	24574182
2008.9		22985700	24854741	24685231

TABLE 2: Comparison of the prediction errors obtained by the three methods for each period.

Method	Test sample period/sample period MAPE (%)	September 2008 forecast MAPE (%)
Neural network method	3.62	6.821
ARIMA method	5.503	0.57
GARCH method	5.582	0.614

7. Discussion

In the context of this paper, the author believes that there are three main ways for China to continue to steadily increase its exports to foreign countries in the context of the global economic crisis.

The government can provide subsidies to increase exports in two ways: firstly, by lowering the costs of enterprises to compensate for the higher export prices and stagnant sales caused by the appreciation of the RMB; secondly, by reducing the chances of losses and bankruptcy of Chinese enterprises through subsidies so that they, especially the powerful ones, can reserve their strength in the economic winter. However, as China has joined the WTO, many policies have to meet the requirements of the WTO, and the excessive use of the previous explicit subsidy policy will lead to many disputes. For example, research subsidies for high-tech enterprises: government investment in basic research projects in universities and government laboratories, followed by public procurement to support the initial application of these results in products and processes, which then spread to commercial applications [16, 17].

The economic crisis has led to a sharp decline in imports, but this is a structural reduction in demand, while demand for high-tech products is still high. The Chinese government should seize this feature and steadily promote the transformation and upgrading of China's export processing trade, encouraging enterprises to produce products with higher technological content, environmental protection and energy saving, and guiding the healthy development of advanced manufacturing and modern production-oriented services so as to identify and meet the new demands of foreign markets at a faster pace [18]. Steady industrial transfer between the east and the west: enterprises in the developed coastal areas of the south-east should transfer production links in the industrial chain to the central and western regions where the factors of production are less expensive and upgrade to

higher value-added links such as design, R&D, and marketing. In this way, the comparative advantage of China's export sector in terms of price can be maintained in the face of the continued appreciation of the RMB [19].

As export trade volumes are the result of a joint game between governments, especially China as a major exporter, they often lead to trade disputes. Therefore, while expanding its imports, China must also focus on a strategic approach and continue to strengthen and improve its international economic and trade relations [20]. For example, China can expand its foreign imports of products in demand at home, for example, by using China's large and risky dollar reserves to buy crude oil, minerals, and raw materials at lower prices and then sell them to the domestic market at lower prices, and by focusing on increasing imports of advanced technology, key equipment, and components [20].

8. Conclusions

The paper first uses six variables as determinants: export value, real exchange rate, Chinese GDP, US IPI, and their seasonal variables, and uses three methods: BP neural network, ARIMA, and AR-GARCH to model China's exports to the US and forecast the next period outside the sample interval. Then, the simulated results of the three models were compared with the real values for the sample period and the forecast period, using the absolute mean percentage error (MAPE) as the error indicator. The results show that all three models are able to simulate and predict China's exports to the US better. The BP neural network model is a good fit for the test period, while the ARIMA and AR-GARCH models have similar results, and they predict the forecast period very well. This is contrary to some papers which state that the neural network method is better than the time series method.

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 author declares that there are no conflicts of interest regarding this work.

References

- [1] S. W. Ho, J. Zhang, and H. Zhou, "Hot money and quantitative easing: the spillover effects of U.S. Monetary policy on the Chinese economy," *Journal of Money, Credit, and Banking*, vol. 50, no. 7, pp. 1543–1569, 2018.
- [2] G. Kapetanios, H. Mumtaz, I. Stevens, and K. Theodoridis, "Assessing the economy-wide effects of quantitative easing," *The Economic Journal*, vol. 122, no. 564, pp. F316–F347, 2012.
- [3] Z. Yang and Y. Zhou, "Quantitative easing and volatility spillovers across countries and asset classes," *Management Science*, vol. 63, no. 2, pp. 333–354, 2017.
- [4] K. Yono, K. Izumi, H. Sakaji, H. Matsushima, and T. Shimada, "Analysis of the macroeconomic uncertainty based on the news-based textual data with financial market," in *Proceedings of the 2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI)*, pp. 661–666, IEEE, Toyama, Japan, 2019, July.
- [5] S. A. Al-Jassar and I. A. Moosa, "The effect of quantitative easing on stock prices: a structural time series approach," *Applied Economics*, vol. 51, no. 17, pp. 1817–1827, 2019.
- [6] K. Yono, K. Izumi, H. Sakaji, T. Shimada, and H. Matsushima, "Measuring the macroeconomic uncertainty based on the news text by supervised LDA for investor's decision making," in *Proceedings of the The International Conference on Decision Economics*, pp. 125–133, Springer, Ávila, Spain, 2019, June.
- [7] Z. Yildirim and M. Ivrendi, "Spillovers of US unconventional monetary policy: quantitative easing, spreads, and international financial markets," *Financial Innovation*, vol. 7, no. 1, pp. 1–38, 2021.
- [8] D. W. Sari, W. Restikasari, S. R. Ajija, H. A. T. Islamiya, and D. Muchtar, "The impacts of foreign direct investment and export expansion on the performance of the high-tech manufacturing industry," *Jurnal Ekonomi Malaysia*, vol. 55, no. 2, pp. 91–105, 2021.
- [9] H. Elsinger, A. Lehar, and M. Summer, "Risk assessment for banking systems," *Management Science*, vol. 52, no. 9, pp. 1301–1314, 2006.
- [10] A. D. Fofack, A. Aker, and H. Rjoub, "Assessing the post-quantitative easing surge in financial flows to developing and emerging market economies," *Journal of Applied Economics*, vol. 23, no. 1, pp. 89–105, 2020.
- [11] G. Cicceri, G. Inserra, and M. Limosani, "A machine learning approach to forecast economic recessions-an Italian case study," *Mathematics*, vol. 8, no. 2, p. 241, 2020.
- [12] M. A. Cole, R. J. R. Elliott, and B. Liu, "The impact of the Wuhan Covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach," *Environmental and Resource Economics*, vol. 76, no. 4, pp. 553–580, 2020.
- [13] W. T. Woo and W. Zhang, "Combating the global financial crisis with aggressive expansionary monetary policy: same medicine, different outcomes in China, the UK and USA," *The World Economy*, vol. 34, no. 5, pp. 667–686, 2011.
- [14] "Center for macroeconomic research of xiamen university. (2014). forecast of china's macroeconomic outlook for 2013-2014," *Current Chinese Economic Report Series*, springer, New York, pp. 9–18, 2013.
- [15] Z. Zhengwan, Z. Chunjiong, L. Hongbing, and X. Tao, "Multipath transmission selection algorithm based on immune connectivity model," *Journal of Computer Applications*, vol. 40, no. 12, p. 3571, 2020.
- [16] A. B. R. Costa, P. C. G. Ferreira, W. P. Gaglianone, O. T. C. Guillén, J. V. Issler, and Y. Lin, "Machine learning and oil price point and density forecasting," *Energy Economics*, vol. 102, Article ID 105494, 2021.
- [17] A. Ansar, B. Flyvbjerg, A. Budzier, and D. Lunn, "Does infrastructure investment lead to economic growth or economic fragility? Evidence from China," *Oxford Review of Economic Policy*, vol. 32, no. 3, pp. 360–390, 2016.
- [18] A. Guariglia, W. Hou, X. Hua, and Y. Huang, "Chinese capital markets: the importance of history for modern development," *The European Journal of Finance*, vol. 24, no. 16, pp. 1369–1374, 2018.

- [19] D. Park, A. Ramayand, and K. Shin, "Capital flows during quantitative easing: experiences of developing countries," *Emerging Markets Finance and Trade*, vol. 52, no. 4, pp. 886–903, 2016.
- [20] E. P. Caldentey, "Quantitative easing (QE), changes in global liquidity, and financial instability," *International Journal of Political Economy*, vol. 46, no. 2-3, pp. 91–112, 2017.

Research Article

An Intelligent Classification Method of Multisource Enterprise Financial Data Based on SAS Model

Xiuyan Xu 

Xijing University, Xi'an, Shaanxi 710123, China

Correspondence should be addressed to Xiuyan Xu; 20140034@xijing.edu.cn

Received 10 January 2022; Revised 10 February 2022; Accepted 15 February 2022; Published 24 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiuyan Xu. 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.

An enterprise is often faced with a large amount of financial information and data information. It is inefficient to rely solely on manual work, and the accuracy is difficult to guarantee. For the multisource data of corporate finance, it is more difficult for financial personnel to accurately analyze the connections between the data. For the multisource financial data of enterprise, this is also a time-consuming and laborious task for financial personnel. At the same time, it is difficult to find the correlation between multiple sources of data and then formulate financial data that guides the development of the enterprise. With the advancement of intelligent algorithms, an intelligent classification algorithm similar to the SAS model has emerged, which can realize the intelligent classification of enterprise financial multisource data and accurately predict the future development trend, which is extremely beneficial to the development and performance of the enterprise. This article mainly combines the financial intelligence classification model SAS with clustering and decision tree methods to classify the financial multisource information and uses the neural network method to carry out the future development trend of corporate finance. The research results show that the maximum error of enterprise financial classification after using the intelligent classification method is only 3.71% and that the forecast error of the future development trend of enterprise finance is only 1.77%. This is an acceptable error range, and this intelligent classification method is also greatly improving the efficiency of corporate financial management.

1. Introduction

The development of an enterprise is not only affected by product performance. As the economic system changes, more companies pay more and more attention to the development and fluctuations of corporate finance, which will reflect to a certain extent the operation of the company in each time period [1]. The status provides a reference for the next development plan. The product level and quality of an enterprise only determine the situation of the product itself, and the finance of the enterprise is an important reference for determining development decisions and price setting. Therefore, the development of an enterprise often requires a large number of financial personnel as the support for the development of the enterprise. For a small company, the efficiency of financial personnel can support the analysis and planning of the company and products, but for medium and large enterprises, this task is often difficult, and the efficiency

of corporate financial personnel and human factors are often limited [2, 3]. At the same time, with the development of economic globalization, the financial data of enterprises is not just a simple single-input and single-output relationship. SAS model is an important classification and prediction model in the field of intelligent finance. Using SAS model can solve the classification and prediction problems of multisource financial data that are difficult to handle for enterprise financial personnel, which is convenient [4].

In the face of economic globalization and the prevailing economic era of e-commerce, the development of enterprises is bound to generate data and information from multiple channels [5]. Corporate financial information is a necessary means to guide the company long-term development and make changes in accordance with economic conditions [6]. For the development of an enterprise, it is often not enough to merely rely on the improvement of product quality and technological level. A large enterprise

often has a large financial team. However, with the development of smart financial economy in recent years, smart algorithms have shown good performance and applicability in the financial field [7]. The intelligent financial classification method will not only improve the classification and processing of financial data, but it will also analyze the future financial trend of the enterprise based on the relationship between multisource data for reference by enterprise decision makers. This method also reduces the complicated business volume of the financial staff and improves the efficiency of the financial staff [8]. For multisource data in finance, it is difficult for financial personnel to discover the nonlinear correlations only by relying on their own knowledge level. Intelligent algorithms will assist humans to recognize these auxiliary nonlinear correlations [9], whether they are the trend of data changes or the relationship with time.

With the diversification of corporate financial data and the continuous advancement of intelligent classification methods, a lot of research on intelligent financial classification methods has been carried out in the field of corporate finance, and many excellent results have been also achieved [10]. Zhang et al. [11] used the firefly algorithm to optimize the support vector machine (SVM) algorithm and developed a new type of firefly support vector machine (FA-SVM) algorithm. The credit risk of supply chain finance was evaluated and verified, and the results showed that the algorithm they proposed has high accuracy in the financial field and can accurately find groups with lower credit. Chen et al. [12] studied the relationship between the Internet and corporate financing efficiency using the data envelopment analysis method and the slack based measure method (DEA-SBM). At the same time, they analyzed the role of corporate financing spillovers by means of spatial measurement and concluded that smart methods can help improve the financing rate of enterprises and accurately capture the spatial spillover rate. In response to the large workload and error-prone problems in current corporate financial revenue forecasting, Huang and Huang [13] proposed a support vector machine-based corporate financial revenue forecasting model, which is based on the method of the average mean square error, absolute error value (MAPE), and other quantitative parameters. The analysis and the conclusion show that the support vector machine method is suitable for enterprise financial revenue forecasting. Wang et al. [14] believe that smart methods can identify target risk rates and reduce risks in corporate crowdfunding tasks. They applied decision tree, logistic regression, support vector machine, and deep learning methods to compare risk identification, and the accuracy rate reached 92.3%. This is a beneficial way for Internet finance. Sang et al. [15] used genetic algorithm and BP neural network to predict and analyze the dynamic changes of financing services and commodity flows of small- and medium-sized enterprises. The accuracy rate of the BP neural network algorithm for the evaluation of the supply chain of small and medium enterprises has reached 89.3%. This is a relatively low accuracy for small and medium enterprises. The loss of bank interest rates makes it possible. Zhu et al. [16] combined two traditional machine learning

methods, random subspace (RS) and MultiBoosting, to propose a hybrid enhanced machine learning model. The results show that this method has good feasibility and accuracy when dealing with small and medium samples in enterprises. Tubastuvi [17] believes that small- and medium-sized enterprises have many problems and that the problem of funding is particularly prominent. He studied the impact of loss-sharing PLS contracts on financing channels and used structural equation model SEM to classify the data. The conclusion shows that a higher PLS contract will promote the development and financing of small- and medium-sized enterprises. Li [18] studied the debt relationship in shipping companies, proposed an intelligent strategy model, and divided shipping companies' data into explicit and invisible forms. The results showed that this method is more efficient than the traditional shipping economic evaluation model. Zdravkovic et al. [19] integrated industrial Internet technology, distributed systems, cloud computing, and deep learning technology to study the application of artificial intelligence technology in enterprises. This method can realize the independent decision-making function of corporate finance.

From the above review of the development status of corporate finance, it can be seen that the current research on corporate finance mainly uses machine learning and other types of algorithms to study the trend of a single data source. This study analyzes the performance of intelligent classification of multisource data of corporate finance, and good intelligent classification makes sense for the business [20]. Intelligent financial classification will reduce the workload of financial staff and reduce the error rate of thinking. At the same time, financial staff will make full use of professional knowledge to make more advanced classification planning [21]. If an enterprise can well combine the advantages of smart finance and manual finance, the company will respond well to the development trend of the global economy, which is beneficial to the long-term development of an enterprise [22]. Moreover, the development of computer hardware and the advancement of intelligent classification algorithms at this stage provide more support for intelligent financial classification, such as the SAS model library [23, 24]. Decision trees, support vector machines, neural networks, etc. are all excellent classification models in the SAS model library, which are easy to build, and perform predictive analysis for intelligent financial classification and forecasting. SAS provides more convenience for the classification of smart finance.

This article is mainly divided into 5 sections to introduce the feasibility of intelligent financial classification. The first section introduces the current situation and necessity of the development of enterprise financial intelligence. The second section mainly introduces the significance of the development of intelligent financial classification for corporate performance and the source of the data set. The third section introduces the main algorithms and processes to realize the intelligent financial classification of enterprise performance. The fourth section explains the feasibility and accuracy of classification methods and forecasting methods in enterprise financial forecasting and classification. The fifth section is a

summary of the feasibility analysis of enterprise financial intelligence classification tasks.

2. The Necessity and Data Source of SAS Model for Intelligent Classification of Enterprise Multisource Data

2.1. The Significance of Smart Financial Data Classification. With the development of economic globalization and the continuous development of the e-commerce economy, the economic system and business model of enterprises have undergone major changes compared to previous years [25]. Enterprise data also come from multiple sources. If only financial personnel are used to process cumbersome financial data, it will be difficult, and it is easy to produce certain wrong information. The SAS model library can provide a variety of intelligent forecasting algorithms for corporate financial analysis, and the construction is simple and efficient, such as decision trees, random forests, and neural network algorithms [26]. The sources of multisource financial data are different, and the characteristics of the data information are also different. Not only is corporate financial staff difficult and cumbersome to deal with, but it is also difficult to find the correlation between its multisource data and to use the data itself to predict the development trend of the corporate economy [27, 28]. Economic development is changing rapidly, which requires business managers to have insight into financial data and information and then make corresponding business development decisions [29]. The SAS model can find the correlation between these multisource data and provide new discoveries and clues for the financial staff [30]. Once the mapping relationship between multisource data is established through SAS, corporate financial personnel can use their own knowledge to further lock the relationship between multisource data and corporate products, which is a very simple way to provide corporate managers with development planning [31]. Intelligent financial data classification will not only improve the efficiency of financial personnel, but also formulate a long-term development plan for enterprises. This is a necessary step for enterprises to maintain long-term development in today economic globalization [32]. The SAS model library is a modeling application library proposed for intelligent finance. It contains functions such as data classification and data prediction. The algorithms basically included the current popular classification and forecasting algorithms, and there are a series of postprocessing function algorithms for data statistics, which is relatively easy for financial personnel to apply and build. The combination of the advantages of the SAS model and multisource data is a meaningful and necessary thing for the intelligent development of corporate finance.

2.2. Method and Source of Data Acquisition. For an enterprise, other financial information such as product information, employee information, and revenue performance generated every day is complex and multisource. In the today era, company financial information comes from

multiple sources, such as the company employee performance information, employee bonus information, product marketing information, and corporate financing information. Not only are these sources of information different, but the proportions occupied and the contribution to the long-term development of an enterprise are also different. How to better fully dig out the relevance of these financial information, make better classification, and make predictions based on the classification information is meaningful for the development of an enterprise. If such a complicated corporate financial data only relies on financial personnel to classify and predict the development trend of the corporate economy, it will be a laborious task, and the accuracy rate will be difficult to guarantee. At the same time, the invisible correlation between these financial data is difficult to discover only by relying on the experience and knowledge of financial personnel. Artificial intelligence methods have been proven in many fields to discover the mapping relationships and correlations between high-dimensional and nonlinear data. For the application of the SAS model, this paper selects the employee performance, the corporate financing, the product sales, and the overall profit and loss of the company as multiple data sources for intelligent classification. Because this type of corporate financial data sources is complex, if the SAS model can better intelligently perform this type of financial data classification, this is easily extended to other areas of corporate finance. Figure 1 shows the process of intelligent classification of enterprise financial multisource data through SAS model. Multisource data mainly includes employee performance, product marketing, corporate financing, and e-commerce marketing performance.

3. Intelligent Classification Method of Enterprise Financial Multisource Data

3.1. The Introduction of SAS Model. The SAS model is a commonly used data analysis model in financial management. It mainly refers to a comprehensive discipline model collectively referred to as financial analysis, data management, corporate economic analysis, and accounting analysis. It can give full play to its own advantages to sort out and summarize the past economic conditions of the enterprise, analyze its status quo, and predict its future development trend. It is also an important analysis model for the evaluation of enterprise profit and loss, investment recovery status, and profit and loss of marketing activities. It can be seen that its application range is relatively wide and it has also gained the trust of enterprises and financial personnel. The financial analysis of a company mainly includes corporate strategy analysis, accounting analysis, and financial statement analysis. The above three types of analysis are called SAS financial analysis. This article aims mainly to intelligently classify multisource financial data of enterprises and make effective predictions. Since the sources of corporate financial data are complex and diverse, these redundant data require intelligent algorithms to perform intelligent classification based on the internal connections of the data themselves. These classification algorithms are

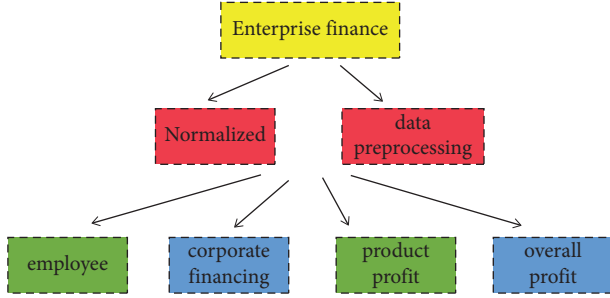


FIGURE 1: The processing flow of enterprise financial multisource data by SAS model.

relatively easy to implement for the SAS model. At the same time, the SAS model can be used to classify the financial data of the enterprise, and then an effective evaluation and prediction can be made. The decision tree and clustering algorithm in the SAS model library have been well proven in the field of corporate finance. They have a good classification effect. The data in this paper is multisource data of corporate finance, which requires the SAS classification model.

3.2. The Introduction of Intelligent Classification Algorithm.

The sources of enterprise financial data are diverse and complex. When only relying on the processing of financial personnel, the work is extremely complicated and prone to errors. At the same time, there are nonlinear high-dimensional correlations between multisource financial data, which makes it difficult for financial personnel to rely solely on professional knowledge. The recent rapid development of intelligent financial technology is a good thing for the development of financial personnel and enterprises. The SAS model is a commonly used model in the financial field. It can not only effectively classify financial multisource data, but also effectively predict the future trend of the enterprise economy, which is meaningful task to enterprise managers. There are many classification methods in the SAS system model, such as decision trees, support vector machines, logistic regression, and clustering. The schematic in Figure 2 shows the application of the clustering method from the SAS model to the intelligent classification of enterprise financial multisource data. It can be seen that clustering can classify four enterprise data sources according to different distances. The sample images were obtained from the website <https://image.so.com>.

Clustering is a type of machine learning method that can classify data based on the distance between data or the density of data. The standard for evaluating the quality of the classification results is that the distances between different types of data are as large as possible, and the distances or the differences between the densities of the same type of data are as small as possible. In this paper, the multisource financial data of enterprises needs to be divided into four categories: employee performance, product quality, corporate financing performance, and enterprise profit and loss. Therefore, the clustering method based on distance is adopted.

Equation (1) shows the expression of the external evaluation index, where a , b , c , and d represents employee

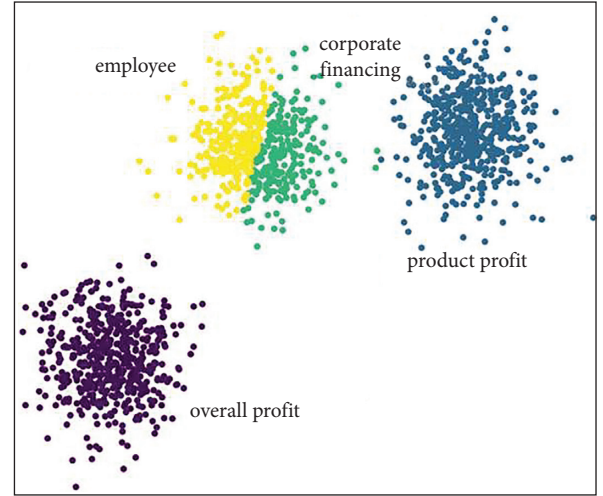


FIGURE 2: Classification of enterprise financial data using clustering methods.

performance, enterprise product profit and loss, enterprise financing, and overall enterprise profit and loss in enterprise financial multisource data.

$$R = \frac{a + d}{a + b + c + d}. \quad (1)$$

Equation (2) shows the index of the Rand statistic, where P represents the precision rate and R represents the recall rate. β is a constant parameter; when $\beta = 1$, this is the most common F1-measure.

$$F = \frac{(\beta^2 + 1)PR}{\beta^2 P + R}. \quad (2)$$

Equation (3) shows the FM parameters of external indicators, where P is the accuracy rate and R is the recall rate. The larger the FM value, the better the clustering effect of corporate finance.

$$FM = \sqrt{\frac{a}{a+b}} * \sqrt{\frac{a}{a+c}} = \sqrt{P * R}. \quad (3)$$

Common measurement methods include Euclidean distance, Manhattan distance, and Chebyshev distance. These indicators evaluate the clustering effect from different angles. It is mainly aimed at the evaluation of samples, which is a more refined method of evaluation. Equation (4) shows the expression for the Euclidean distance, a measure of the distance between two points in space.

$$d_{i c t_e d} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2}. \quad (4)$$

Equation (5) shows the Chebyshev distance, an evaluation index of clustering effect, which defines a point in space as the difference between the absolute values of the coordinate values.

$$dict_{cd} = \lim_{t \rightarrow \infty} \left(\sum_{k=1}^m |x_{jk} - y_{ik}|^t \right)^{1/t}. \quad (5)$$

Equation (6) shows the Minkowski distance evaluation index, which is a measure in Euclidean space, and also a variant extension of Euclidean distance and Manhattan distance. When P takes different values, it represents different distance evaluation criteria.

$$dict_{mind} = \sqrt[p]{\sum_{k=1}^m |x_{ik} - y_{jk}|^p}. \quad (6)$$

3.3. Time Correlation Algorithm of Enterprise Financial Multisource Data. After clustering, the multisource financial data of enterprises can be divided into four categories: employee performance, product quality, corporate financing performance, and overall corporate profit and loss. For an enterprise manager, this is not only concerned with the classification of enterprise multisource data. The future of economic development trend of enterprises is also the focus of attention for the business manager. Combining the characteristics of enterprise data itself, this paper adopts long short-term memory recurrent neural network to predict the future development trend of enterprise financial data. The advantage of long short-term memory neural network from convolutional neural network is that it can effectively map the temporal correlation between data. However, there is not only the correlation between categories, but also a certain time correlation between the data itself. The development of an enterprise is often related to policies in different time periods and the development trend of the global economy. Therefore, the long short-term memory recurrent neural network is used to predict the temporal correlation between the multisource financial data of the enterprise and intuitively show the prediction results to the enterprise decision makers, which will assist the enterprise to make decisions. Figure 3 shows the prediction process of multisource financial data of enterprises.

Equation (7) shows the expression of the “forget” gate of the long short-term memory neural network, which selectively accepts part of the input at the previous moment and the input at this moment and combines them as the input value of the next gate structure. ω is the weights, s_b^{t-1} is the state value of the last moment $t-1$, and x_i^t is the input parameter.

$$\alpha_i^t = \sum_{i=1}^k \omega_{ii} x_i^t + \sum_{b=1}^v \omega_{ct} s_b^{t-1}. \quad (7)$$

Equations (8) and (9) show the structure of the input gate. This layer mainly performs feature extraction for the input of the “forget” gate, and it undergoes nonlinear transformation through the activation function.

$$b_g^t = f(\alpha_g^t), f = \tanh, \quad (8)$$

$$\vec{C}_t = \tanh(w_c \bullet [h_{t-1}, P_t] + b_c). \quad (9)$$

It can refresh the variable by the following formula:

$$\vec{b}_t = f_t \times \vec{b}_{t-1} + i_t \times \vec{b}_t. \quad (10)$$

Equations (11) and (12) show the output gate structure of the neural network structure, which is a layer for final feature input that selectively inputs temporal features and combines the state information of different historical moments and the feature information of the input gate.

$$s_c^t = b_g^t s_c^{t-1} + b_i^t g(a_c^t), \quad (11)$$

$$\vec{h}_t = O_t \times \tanh(\vec{C}_t). \quad (12)$$

3.4. The Preprocessing of Multisource Financial Data of Enterprise. The data preprocessing stage is an important stage for using the SAS model. This is because the sources of corporate financial data are different and there are certain gaps in characteristics. The purpose of data preprocessing is to normalize different financial data sources so that these data conform to the same distribution and are in the same order of magnitude, which is beneficial for both classification algorithms and enterprise time series predictions. If these multisource financial data are not preprocessed, this will cause uneven weight distribution in the classification and prediction process, which will not only reduce the learning ability of the model but also reduce its generalization ability. In other words, this model may show better accuracy only for specific corporate finances and will produce poor results when forecasting and categorizing the financial multisource data of other companies. From the above description, it can be learned that the preprocessing of multisource financial data is not only an important step of the SAS model, but also an important step to improve the accuracy and generalization ability of the model. In this study, multisource financial data will be processed into distribution characteristics that conform to the normal distribution, and the data will be processed into an order of magnitude between 0 and 1. The standard processing methods are adopted in this paper to preprocess multisource enterprise data.

4. Classification and Predictive Analysis of Enterprise Financial Multisource Data

This research first analyzes the results of intelligent classification of the enterprise’s multisource financial data. It can be clearly seen from Figure 4 that the errors of the classification results are all within 4%, which is a trustworthy result for enterprise managers to make decisions. This article mainly classifies employee performance, product marketing, corporate financing rate, and overall profit and loss. It can be seen that the smallest classification error is only 1.77%, which is an acceptable error for a large enterprise. This part of the error is mainly from employee performance. The reason why this part is the smallest is mainly that the performance of employees is within a controllable range. The largest error is 3.77%. This part of the error comes from the

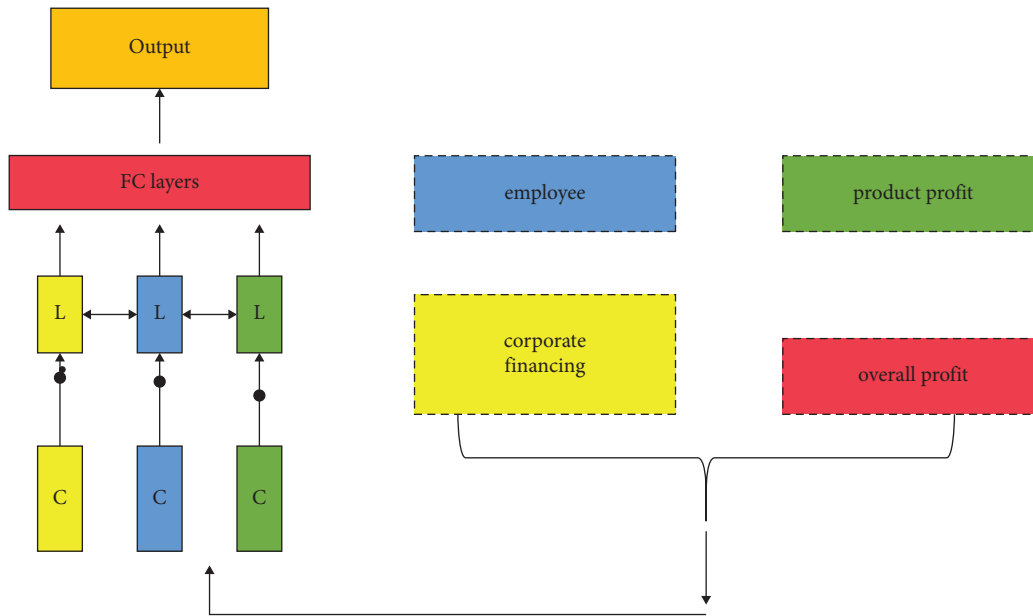


FIGURE 3: Prediction process of multisource financial data of enterprises.

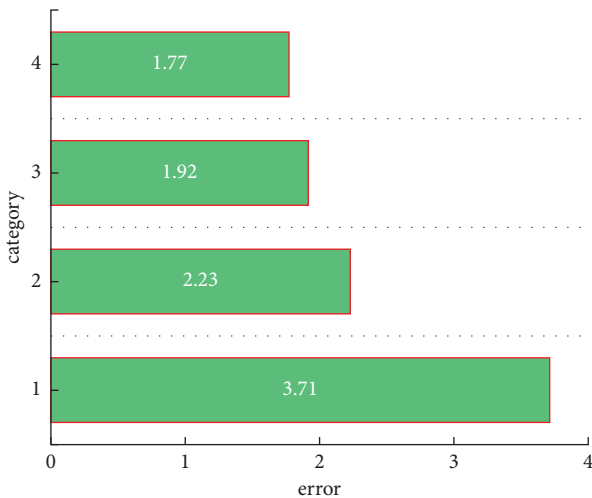


FIGURE 4: The classification error of multisource financial data.

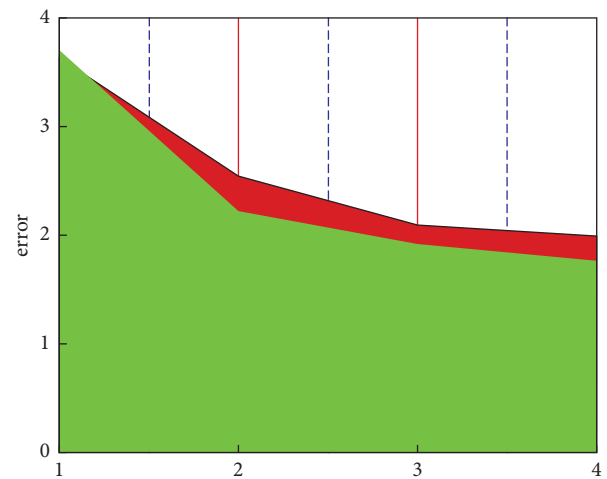


FIGURE 5: Forecast and actual development trend of corporate financial data.

overall profit and loss of the enterprise. The reason for this part of the error is that the overall profit and loss of the enterprise is a complex and changeable variable, and there are certain differences between it and these multisource data. Part of the correlation also comes from accumulated errors. Generally speaking, this model is credible in the intelligent classification of enterprise financial multisource data. Figure 5 reflects the forecast and actual development trend of multisource financial data of enterprises. For these four types of classified multisource data, the overall forecast trend is consistent with the actual development trend of the company, but there is a small error. From a macro perspective, the difference between the development trend of corporate finance and the multisource data is better mapped, which shows that this model has certain accuracy in the prediction of corporate multisource finance. For the

prediction of employee performance, there is basically no error, but for the other three predictions, there are certain errors. This is mainly due to the sudden change of corporate finance over time.

In order to further analyze the accuracy and feasibility of the SAS model in the classification and forecasting of corporate financial data, Figure 6 shows a box diagram of the forecasted value of financial multisource data and the development trend of actual data. It can clearly indicate not only the trend and accuracy of the predicted value, but also the overall distribution trend of the predicted value and the actual value. From the overall point of view in Figure 6, the predicted value of corporate financial data is generally higher than the actual data, but the error between them is relatively small. At the same time, the predicted value of financial data

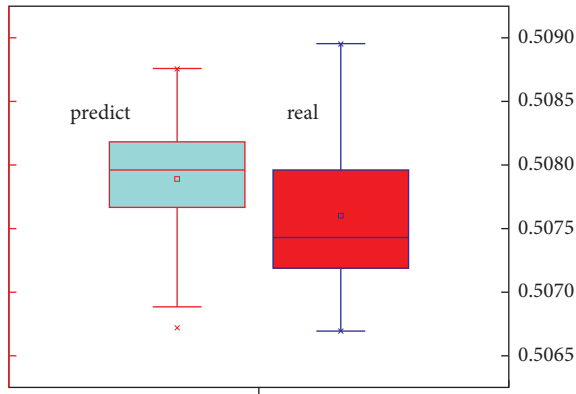


FIGURE 6: Box diagram of forecast and real value of enterprise financial multisource data.

is relatively more concentrated than the actual financial data, which is mainly caused by the concentrated distribution of weights. It can be seen from Figure 6 that the predicted value of the multisource data of corporate finance and the actual value are basically between 0.506 and 0.509, and the error is relatively small. Figure 7 shows the linear correlation between the predicted value of corporate financial multisource data and the actual data, which can show the stability of the predicted value. It can be seen from Figure 7 that the data points are distributed on both sides of the linear function and that the distance from the straight line of the linear function is relatively small, which shows that the forecast of multisource financial data maps the actual data value well. It can also be seen from the figure that only a small part of the data deviates from the linear function and the distance is relatively large. This may be due to the large randomness of this type of data source and the obvious relationship with time, which requires the increase of this part of the data to improve the accuracy of this part of the forecast and classification.

Figure 8 shows the normal distribution of errors for multisource corporate financial data. It can be seen that the distribution of multisource financial data errors is a relatively average one, and the data points are all within the upper and lower boundaries of the normal distribution. Only some of the financial data points are off the center line, but the error in this part of the data is also within an acceptable range. The correlation coefficients all exceed 0.96, which shows that the financial data of enterprises has a good prediction effect. In addition, the data points are basically distributed on both sides of the $y = x$ function, which means that there is a good correlation. Figure 9 shows the forecast classification error trends for the two types of financial data sources. It can be seen that the errors fit the normal distribution curve well, which is the same as the input financial data source in the preprocessing process. Moreover, the distribution of the above two types of financial data with large errors is relatively uniform, which means that the SAS model is suitable for the classification and prediction of multisource financial data of enterprises.

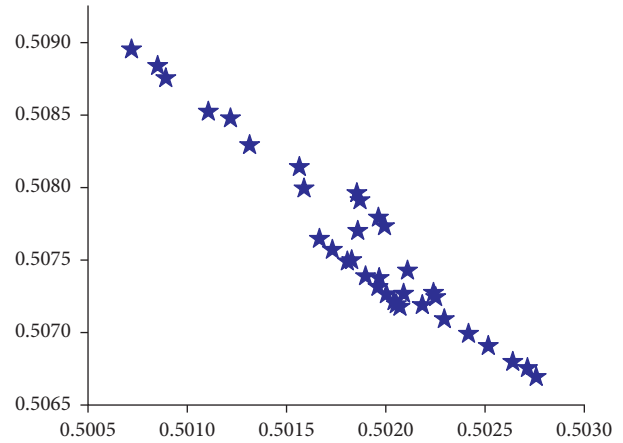


FIGURE 7: Linear correlation of multisource data of enterprise finance.

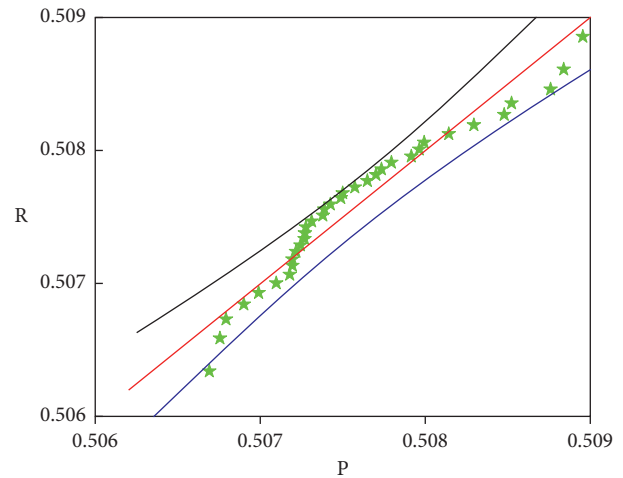


FIGURE 8: Predictive distribution of corporate financial data.

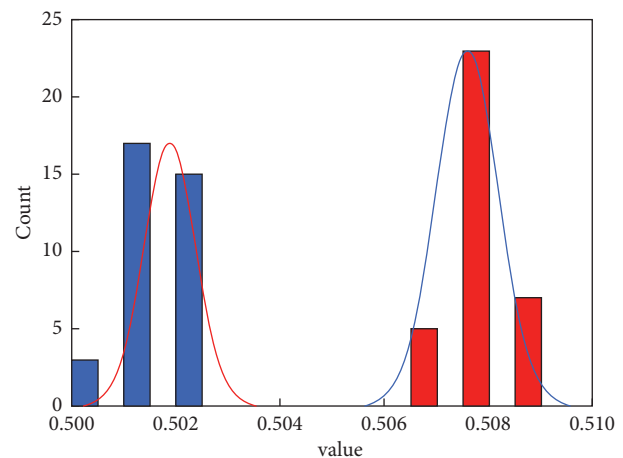


FIGURE 9: Cumulative error distributions for two types of financial data.

5. Summary of Research on Intelligent Classification of Enterprise Financial Data

The sources of corporate financial data are often diverse, complex, and changeable due to the development trend of economic globalization and the continuous stimulation of the e-commerce economy. It would be a laborious and error-prone task to rely solely on financial staff to organize and forecast a company's multisource financial data. SAS model is a model specially used for enterprise financial analysis and prediction. It can quickly and accurately realize intelligent classification of financial multisource data, and the efficiency is high.

This paper firstly uses clustering to effectively classify employee performance, product profit and loss, corporate financing, and overall profit and loss and uses the classification results to accurately predict the financial development trend of the company. In general, the classification accuracy is relatively high, the largest error is only 3.71%, and the smallest error is only 1.77%. The error is largest compared to other factors, which is due to the large randomness of the overall profit and enterprise loss, and it is also affected by the changes of socioeconomy. The other three types of financial data have cumulative errors. After a reasonable classification by the SAS model, these classified data are used to predict the future development trend of the enterprise finance. Generally speaking, the financial forecast of the enterprise is in good agreement with the development trend and change trend of the actual data value, which is a trustworthy model for the enterprise. The linear correlation between them is also relatively good. The data points are distributed on both sides of the linear function, and the distance from the linear function is relatively close, which shows that this model is feasible in the classification and prediction of enterprise financial intelligence.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The author thanks the Scientific Research Program Funded by Shaanxi Provincial Education Department, "Research on the Influence of Internal Audit Quality on the Investment Efficiency of Enterprises" (no. 21JK0409).

References

- [1] T. Renault, "Sentiment analysis and machine learning in finance: a comparison of methods and models on one million messages," *Digital Finance*, vol. 2, no. 1, pp. 1–13, 2020.
- [2] B. J. N. Heaton, G. N. H. J., and P. Witte, "Deep learning for finance: deep portfolios," *Applied Stochastic Models in Business and Industry*, vol. 33, no. 1, pp. 3–12, 2017.
- [3] L. Ic and Y. Su, S. Js, "Machine learning for enterprises: applications algorithm selection, and challenges," *Business Horizons*, vol. 63, no. 2, pp. 157–170, 2020.
- [4] J. Sirignano and C. Rc, "Universal features of price formation in financial markets: perspectives from deep learning," *Quantitative Finance*, vol. 19, no. 9, pp. 1449–1459, 2019.
- [5] A. Korbi and L. Lleshaj, "Finance leasing and arma forecasting: evidence from Albania," *European Journal of Business and Management Research*, vol. 5, no. 6, pp. 10–18, 2020.
- [6] Z. Jian, Q. Li, and L. Tc, "Market revenue prediction and error analysis of products based on fuzzy logic and artificial intelligence algorithms," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–8, 2020.
- [7] G. Baryannis, S. Dani, and G. Antoniou, "Predicting supply chain risks using machine learning: the trade-off between performance and interpretability," *Future Generation Computer Systems*, vol. 101, pp. 993–1004, 2019.
- [8] M. Kim, J. Jeong, and S. Bae, "Demand forecasting based on machine learning for mass customization in smart manufacturing," in *Proceedings of the 2019 International Conference on Data Mining and Machine Learning*, pp. 6–11, New YorkNY, April 2019.
- [9] M. Karner, R. Glawar, W. Sihn, and K. Matyas, "An industry-oriented approach for machine condition-based production scheduling," *Procedia CIRP*, vol. 81, pp. 938–943, 2019.
- [10] A. Kara and I. Dogan, "Reinforcement learning approaches for specifying ordering policies of perishable inventory systems," *Expert Systems with Applications*, vol. 91, pp. 150–158, 2018.
- [11] H. Zhang, Y. Shi, X. Yang, and R. Zhou, "A firefly algorithm modified support vector machine for the credit risk assessment of supply chain finance," *Research in International Business and Finance*, vol. 58, Article ID 101482, 2021.
- [12] Q. Chen, Y. Zhang, and L. Chen, "A study of Internet development and enterprise financing in China," *Networks and Spatial Economics*, vol. 21, no. 3, pp. 495–511, 2021.
- [13] L. Huang and C. L. Huang, "Comparison of multiple machine learning models based on enterprise revenue forecasting," in *Proceedings of the 2021 asia-pacific conference on communications technology and computer science*, vol. 1, pp. 354–359, Shenyang, China, January 2021.
- [14] W. Wang, H. Zheng, and Y. J. Wu, "Prediction of fundraising outcomes for crowdfunding projects based on deep learning: a multimodel comparative study," *Soft Computing*, vol. 24, no. 11, pp. 8323–8341, 2020.
- [15] B. Sang, "Application of genetic algorithm and BP neural network in supply chain finance under information sharing," *Journal of Computational and Applied Mathematics*, vol. 384, Article ID 113170, 2021.
- [16] Y. Zhu, L. Zhou, C. Xie, G.-J. Wang, and T. V. Nguyen, "Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach," *International Journal of Production Economics*, vol. 211, pp. 22–33, 2019.
- [17] N. Tubastuvi, "The role of profit-and-loss sharing contracts in strengthening financing access of small medium enterprise's (SME's): the case of Indonesia," *Advanced Science Letters*, vol. 24, no. 1, pp. 129–132, 2018.
- [18] X. Li, "Debt financing cost evaluation method of shipping enterprises," *Journal of Coastal Research*, vol. 103, no. sp1, pp. 744–748, 2020.
- [19] M. Zdravkovic, H. Panetto, and G. Weichhart, "AI-enabled enterprise information systems for manufacturing," *Enterprise Information Systems*, Article ID 1941275, 2021.

- [20] P. Hajek and M. Z. Abedin, "A profit function-maximizing inventory backorder prediction system using big data analytics," *IEEE Access*, vol. 8, pp. 58982–58994, 2020.
- [21] A. Kuhnle, L. Schäfer, N. Stricker, and G. Lanza, "Design, implementation and evaluation of reinforcement learning for an adaptive order dispatching in job shop manufacturing systems," *Procedia CIRP*, vol. 81, pp. 234–239, 2019.
- [22] R. Masocha, "Does environmental sustainability impact innovation, ecological and social measures of firm performance of SMEs? Evidence from South Africa," *Sustainability*, vol. 10, no. 11, p. 3855, 2018.
- [23] C. Luo, S. Fan, and Q. Zhang, "Investigating the influence of green credit on operational efficiency and financial performance based on hybrid econometric models," *International Journal of Financial Studies*, vol. 5, pp. 10–27, 2017.
- [24] P. Falcone and E. Sica, "Assessing the opportunities and challenges of green finance in Italy: an analysis of the biomass production sector," *Sustainability*, vol. 11, no. 2, p. 517, 2019.
- [25] X. Xu and J. Li, "Asymmetric impacts of the policy and development of green credit on the debt financing cost and maturity of different types of enterprises in China," *Journal of Cleaner Production*, vol. 264, Article ID 121574, 2020.
- [26] J. Abbas, "Impact of total quality management on corporate green performance through the mediating role of corporate social responsibility," *Journal of Cleaner Production*, vol. 242, Article ID 118458, 2020.
- [27] E. Wang, X. Liu, J. Wu, and D. Cai, "Green credit, debt maturity, and corporate investment-evidence from China," *Sustainability*, vol. 11, no. 3, p. 583, 2019.
- [28] J. T. Zhang, M. Zhang, and Z. Q. Guo, "Analysis of automobile supply chain financial risk based on modified KMV model [J]," *Industrial Engineering & Management*, vol. 24, no. 1, pp. 128–135, 2019.
- [29] M.-L. Tseng, K.-J. Wu, J. Hu, and C.-H. Wang, "Decision-making model for sustainable supply chain finance under uncertainties," *International Journal of Production Economics*, vol. 205, pp. 30–36, 2018.
- [30] Á. M. Miguel and A. M. José, "Determinants of the propensity for innovation among entrepreneurs in the tourism industry," *Sustainability*, vol. 12, p. 5003, 2020.
- [31] Y. Shahab, Y. Riaz, C. G. Ntim, Z. Ye, and Q. Zhang, "Online feedback and crowdfunding finance in China," *International Journal of Financial Economics*, pp. 1–19, 2020.
- [32] D. Liu, S. Chen, and S. Chen, "The effect of seasonal equity offerings and external financial dependence on corporate innovation of listed manufacturing companies," *Chin. J. Manag.* vol. 8, pp. 1168–1178, 2019.

Research Article

5G Network Slicing: Methods to Support Blockchain and Reinforcement Learning

Juan Hu ¹ and Jianwei Wu²

¹School of Intelligent Engineering, Zhengzhou University of Aeronautics, Zhengzhou 450046, China

²The 27th Research Institute of China Electronics Technology Group Corporation, Weishi 450047, China

Correspondence should be addressed to Juan Hu; hujian@zua.edu.cn

Received 25 January 2022; Revised 21 February 2022; Accepted 2 March 2022; Published 24 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Juan Hu and Jianwei Wu. 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 advent of the 5G era, due to the limited network resources and methods before, it cannot be guaranteed that all services can be carried out. In the 5G era, network services are not limited to mobile phones and computers but support the normal operation of equipment in all walks of life. There are more and more scenarios and more and more complex scenarios, and more convenient and fast methods are needed to assist network services. In order to better perform network offloading of the business, make the business more refined, and assist the better development of 5G network technology, this article proposes 5G network slicing: methods to support blockchain and reinforcement learning, aiming to improve the efficiency of network services. The research results of the article show the following: (1) In the model testing stage, the research results on the variation of the delay with the number of slices show that the delay increases with the increase of the number of slices, but the blockchain + reinforcement learning method has the lowest delay. The minimum delay can be maintained. When the number of slices is 3, the delay is 155 ms. (2) The comparison of the latency of different types of slices shows that the latency of 5G network slicing is lower than that of 4G, 3G, and 2G network slicing, and the minimum latency of 5G network slicing using blockchain and reinforcement learning is only 15 ms. (3) In the detection of system reliability, reliability decreases as the number of users increases because reliability is related to time delay. The greater the transmission delay, the lower the reliability. The reliability of supporting blockchain + reinforcement learning method is the highest, with a reliability of 0.95. (4) Through the resource utilization experiment of different slices, it can be known that the method of blockchain + reinforcement learning has the highest resource utilization. The resource utilization rate of the four slices under the blockchain + reinforcement learning method is all above 0.8 and the highest is 1. (5) Through the simulation test of the experiment, the results show that the average receiving throughput of video stream 1 is higher than that of video stream 2, IOT devices and mobile devices, and the average cumulative receiving throughput under the blockchain + reinforcement learning method. The highest is 1450 kbps. The average QOE of video stream 1 is higher than that of video stream 2, IOT devices and mobile devices, and the average QOE is the highest under the blockchain + reinforcement learning method, reaching 0.83.

1. Introduction

Relieving users' network congestion, reducing network latency, and offloading the network are the top priorities for 5G networks. As a core technology, the 5G network slicing technology can effectively solve the challenges of business creation and exclusive network access for different users, as well as the coexistence of multiple application scenarios. The 5G network is expected to meet the different needs of users

[1]. 5G network slicing may be a natural solution [2]. A wide range of services required for vertical specific use cases can be accommodated simultaneously on the public network infrastructure. 5G mobile networks are expected to meet flexible demands [3]. Therefore, network resources can be dynamically allocated according to demand. Network slicing technology is the core part of 5G network [4]. The definition of 5G network slicing creates a broad field for communication service innovation [5]. The vertical market targeted by

5G networks supports multiple network slices on general and programmable infrastructure [6]. The meaning of network slicing is to divide the physical network into two virtual networks so that they can be flexibly applied to different network scenarios. The future 5G network will also change the mobile network ecosystem [7]. The 5G mobile network is expected to meet the diversified needs of a variety of commercial services [8]. 5G mobile networks must support a large number of different service types [9]. Network slicing allows programmable network instances to be provided to meet the different needs of users. Blockchain can establish a secure and decentralized resource sharing environment [10]. Blockchain is a distributed open ledger [11] and is used to record transactions between multiple computers. Reinforcement learning algorithms can effectively solve large state spaces [12]. Reinforcement learning is mainly used to solve simple learning tasks [13]. 5G networks are designed to support many vertical industries with different performance requirements [14]. Network slicing is considered an important factor in enhancing the network and has the necessary flexibility to achieve this goal. Network slicing is considered one of the key technologies of 5G network [15]. You can create virtual networks and provide customized services on demand.

2. Overview of Related Theories

2.1. 5G Network Slicing

2.1.1. Network Slicing. Network slicing refers to the implementation of offload management of the network when the network is congested and complicated [16]. When facing the different needs of different users, the network is divided into many pieces to meet customer needs. Moreover, it provides targeted services and assistance.

2.1.2. Network Slice Classification. The ultimate goal of 5G network slicing is to organically combine multiple network resource systems to form a complete network that can serve different types of users. Network slices can be divided into independent slices and shared slices as shown in Table 1:

2.1.3. 5G Network Application Scenarios. The application scenarios of 5G networks are divided into three categories: mobile broadband, massive Internet of Things, and mission-critical Internet of Things [17]. The details are shown in Table 2:

2.2. Blockchain

2.2.1. Definition of Blockchain. The blockchain consists of a shared, fault-tolerant distributed database, and a multi-node network [18].

2.2.2. Blockchain Structure. The block chain is composed of a block header and a block body, which forms into a chain structure through the hash of the parent block [19]. The structure is shown in Figure 1:

The structure contains the parent block hash, timestamp, random number, difficulty, and the Merkle root [20]. Its functions are shown in Table 3:

2.2.3. Blockchain Properties. Blockchain technology has three attributes of distribution, security, and robustness [21], as shown in Table 4:

2.3. Reinforcement Learning

2.3.1. Definition of Reinforcement Learning. Reinforcement learning is one of the methods of machine learning. It is mainly to solve the method of how the agent takes different actions in the environment in order to maximize the accumulated rewards obtained.

2.3.2. Reinforcement Learning Process. In the process of reinforcement learning, the agent needs to make decisions on the information in the environment [22]. At the same time, the environment will also reward the agent for the corresponding behavior, and the agent will enter a new state after the behavior. The process is shown in Figure 2:

2.4. Model Design. The 5G network slicing architecture is composed of network slicing demander, slice management (business design, instance orchestration, operation management), slice selection function, and virtualization management orchestration.

The process of the 5G network slicing model is as follows: network services enter the slice manager through the network slice demander, and the slice manager includes business design, instance arrangement, and operation management. After the slice manager enters the slice selection function, it is divided into shared slice function and independent slice special function, and it can also enter the virtualization management orchestration as shown in Figure 3:

3. Formula

3.1. Blockchain

3.1.1. Scalability within Shards. In the process of verifying the block consensus, the scalability within the shard [23] is as follows:

$$\Phi(B_F, T_F) = \frac{|(B_F - B_H)/b_F|}{T_F} \quad (1)$$

Among them, b_I is the average transaction size, B_{Th} is the block header size, and K is the number of shards.

3.1.2. Scalability of Directory Fragmentation. Assuming that the average transaction size is b_F , the block header size is B_H , and the scalability of the directory fragmentation is as follows:

TABLE 1: Network slice classification.

Independent slice	A slice with a logically independent and complete network function. The slice includes a user data plane, a network control plane, and various user business function films, which can provide a logically independent end-to-end private network service for a specific user group. If necessary, only part of the services of specific functions can be provided.
Shared slice	A shared slice is a specific network slice whose network resources can be used by different independent slices. The slice can provide end-to-end services, and when necessary, it can also only provide partial sharing functions.

TABLE 2: 5G network application scenarios.

5G application scenarios	Application examples	Need
Mobile broadband	4K/8K Ultra HD video, holographic technology, augmented reality/virtual reality	High capacity, video storage
Mass Internet of Things	Massive sensors (deployed in measurement, construction, agriculture, logistics, smart cities, homes, etc.)	Large-scale connection (200,000/km ²), mostly stationary
Mission Critical Internet of Things	Autonomous driving, automated factories, smart grids, etc.	Low latency, high reliability

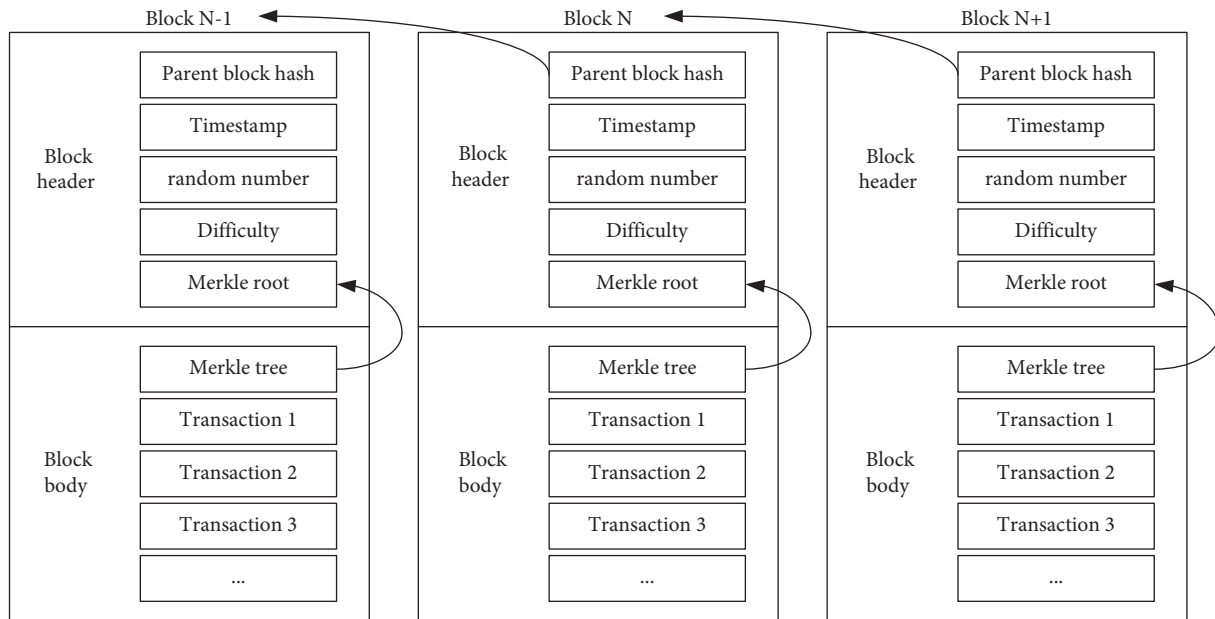


FIGURE 1: Blockchain structure.

TABLE 3: Blockchain structure function.

Parent block hash	Has the hash value of the previous block
Timestamp	Used to record the time when the current block was generated
Random number and difficulty	Used to calculate proof of work
Version number	Record the version of the current block, so that you can view and update the version at any time
Merkle root value of the entire block	It is the root value of the Merkle tree, which stores the hash value of all transactions in the entire block

TABLE 4: Blockchain attributes.

Distributed	The blockchain connects the participating nodes through a peer-to-peer network to realize resource sharing and task allocation between peer nodes. Each network node does not need to rely on the central node and can directly share and exchange information. Each peer node can not only be an acquirer of services, resources, and information but can also be a provider thereof, which reduces the complexity of networking while improving the fault tolerance of the network.
-------------	---

TABLE 4: Continued.

Safety	Blockchain can use encryption technology to asymmetrically encrypt the transmitted data information. The task request for writing data in the blockchain needs to be accompanied by the private key signature of the task initiator. The changed signature is broadcasted together with the task request among participating nodes in the network. Each node can verify its identity, so the task request is not allowed for forgery and tampering. At the same time, the blockchain data structure in the blockchain further ensures that the content in the block cannot be tampered with at will. Even if some nodes in the chain are maliciously forged, tampered with, or destroyed. It will not affect the normal operation of the entire blockchain.
Robustness	The consensus mechanism determines the degree of agreement between the voting weight and computing power between subjects. The entire blockchain system uses a special incentive mechanism to attract more miners to participate in the process of generating and verifying data blocks, perform mathematical calculations in a distributed system structure, use consensus algorithms to select a node, and then create a new one. The effective block of is added to the entire blockchain, and the entire process does not rely on a third-party trusted institution.

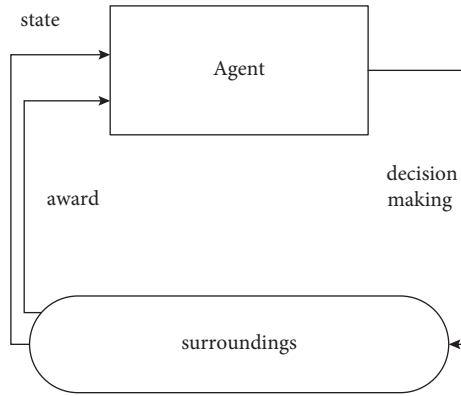


FIGURE 2: Reinforcement learning process.

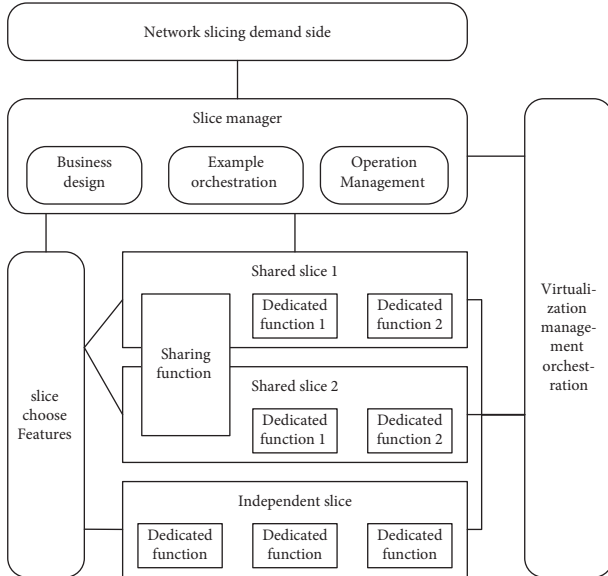


FIGURE 3: Model design.

$$\Phi(B_I, T_{II}) = \sum_{i=1}^K \frac{|(B_I - B_{Ih})/b_I|}{T_{II_i}} \quad (2)$$

3.1.3. Scalability of Sharded Blockchain. The scalability of the entire sharded blockchain is composed of the internal scalability of the shards and the scalability of the catalog shards

[24]. Assuming that the block packing time within the fragment and the directory fragment is the same as T'_I and the block header size is the same as B'_H , the formula is as follows:

$$\begin{aligned} \Phi(B, T) &= \Phi(B_I, T_{II}) + \Phi(B_F, T_F) \\ \Phi(B, T) &= \sum_{i=1}^K \frac{|(B_I - B_{Ih})/b_I|}{T_{II_i}} + \frac{|(B_F - B_h)/b_F|}{T_F} \quad (3) \\ \Phi(B, T) &= \frac{|k(B_I - B'_H)/b_I + (B_F - B_h)/b_F|}{T'_I} \end{aligned}$$

3.2. Reinforcement Learning Methods

3.2.1. Value Function Method. The value function method is to give an estimate of the value for different states. 0 is the given value, and $V^\pi(s)$ starts from state $V^\pi(s)$. The formula is as follows:

$$\begin{aligned} V^\pi(s) &= E_\pi[R|s, \pi], \\ V^\pi(s) &= E_\pi[R_0 + \gamma R_1 + \gamma^2 R_2 + \dots + \gamma^t R_t + \dots | s = S_t]. \quad (4) \end{aligned}$$

The optimal strategy π^* has a corresponding state-value function $V^*(s)$, which is expressed as follows:

$$V^*(s) = \max_{\pi} V^\pi(s) \forall s \in S. \quad (5)$$

In the RL setting, it is difficult to obtain the state transition function P . So, a state-action value function is constructed.

$$\begin{aligned}
Q_\pi(s, a) &= E[R|s, a, \pi], \\
Q^\pi(s, a) &= E[R_0 + \gamma R_1 + \gamma^2 R_2 + \dots + \gamma^t R_t + \dots | s = S_t, a = A_t].
\end{aligned} \tag{6}$$

Given $Q^\pi(s, a)$, in each state, the optimal strategy $\text{argmax}_a Q^\pi(s, a)$ can be adopted. Under this strategy, $V^\pi(s)$ can be defined by maximizing $Q^\pi(s, a)$ as follows:

$$V^\pi(s) = \max_a Q^\pi(s, a). \tag{7}$$

At present, mature deep learning methods such as SARSA and offline Q learning can all be used to solve the value function.

SARSA:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]. \tag{8}$$

Offline Q learning:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)]. \tag{9}$$

3.2.2. Strategy Method. The strategy method is to directly output the action by searching for the optimal strategy π^* . The objective function $J(\theta)$ is defined as the cumulative expected reward.

$$J(\theta) = E \left[\sum_{t \rightarrow 0}^{\infty} \gamma^t r_t | \pi \right] = \int_s d^\pi(S) \int_A \pi(a|s, \theta) R(s, a) ds da. \tag{10}$$

The policy parameter $\nabla_\theta J(\theta)$ is estimated in the discounted cumulative expected reward gradient θ and obtained based on a certain learning rate (α_t). The formula of the strategy gradient method is as follows:

$$\theta_{t+1} = \theta_t + \alpha_t \nabla_\theta J(\theta) |_{\theta=\theta_t}. \tag{11}$$

3.2.3. MDP. MDP mainly solves the problem of learning-related experiences in the interaction between the agent and the environment to achieve the goal [25]. Assuming that the state space is S , it is defined as follows:

$$S = \{(h, x, d, \varphi) | \forall h \in H, x \in X, d \in D, \varphi \in \psi\}. \tag{12}$$

Among them, h represents the state of all wireless channels in the 5G network slice, H represents the channel state space, and H is represented as follows:

$$H = \{(h_1, h_2, \dots, h_m) | \forall m \in M, h_m \in H_m\}. \tag{13}$$

Among them, h_m represents the channel state and H_m represents the channel state space.

x means connection status, X means connection status space. X is defined as follows:

$$X = \{(x_1, x_2, \dots, x_{|M|}) | \forall m \in M, x_m \in X_m\}. \tag{14}$$

d represents the state of all data transmission rates in the slice, and D represents the data transmission rate state space. D is defined as follows:

$$D = \{(d_1, d_2, \dots, d_{|U|}) | \forall u \in U, d \in [d_{\min}, d_{\max}]\}. \tag{15}$$

φ represents the topological state of the physical network, and ψ represents the topological state space in the physical network. ψ is defined as follows:

$$\psi = \{(\varphi_1, \varphi_2, \dots, \varphi_{|N|}) | \forall \varphi \in \psi, \varphi \in \{0, 1\}\}. \tag{16}$$

A_r means that the action space is allocated for unlimited resources, which is defined as follows:

$$A_r = \{(a_{r,1}, a_{r,2}, \dots, a_{r,|U|}) | \forall u \in U, a_{r,u} \in A_{r,u}\}. \tag{17}$$

Among them, $a_{r,u}$ is the 5G network radio resource allocation action, and $A_{r,u}$ is its corresponding network action space, expressed as follows:

$$A_{r,u} = \{v'_{u,m} | m \in M, v'_{u,m} \in [0, v'_m]\}. \tag{18}$$

Among them, $v'_{u,m}$ represents occupied wireless resources.

3.3. Model Building. Suppose the weighted undirected graph of the physical network is $C = (A_i, S_i)$, where the set of network nodes is denoted as $A_i = \{a_1, a_2, \dots, a_n\}$, the calculation level of A_i is denoted as $S_i = \{s_1, s_2, \dots, s_n\}$, and the link set composed of nodes is denoted as $L_n = \{l_1, l_2, \dots, l_n\}$.

The first dynamic dispatch queue state transition function is as follows:

$$Z(X) = \gamma_i + (A_i S_i - L_n) d_i. \tag{19}$$

The second dynamic scheduling queue state transition function is as follows:

$$Z(Y) = \frac{\gamma_i + A_i L_n d_i}{Z(X)}. \tag{20}$$

Combining the above analysis, the 5G network slicing model, the formula is expressed as below:

$$F(x) = \frac{Z(X) + Z(Y)}{d_i} \gamma_i. \tag{21}$$

4. Experiment

4.1. Model Test

4.1.1. Variation of Time Delay with the Number of Slices. This article mainly studies 5G network slicing methods to support blockchain and reinforcement learning. First, we will test the model and compare the blockchain + reinforcement learning method with the blockchain, reinforcement learning, and unused methods. The results are shown in Figure 4.

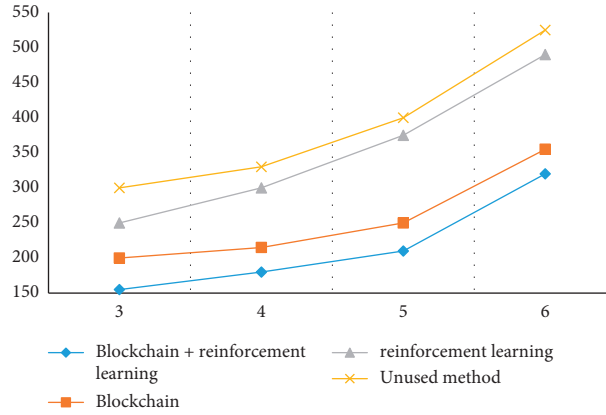


FIGURE 4: Time delay changes with the number of slices.

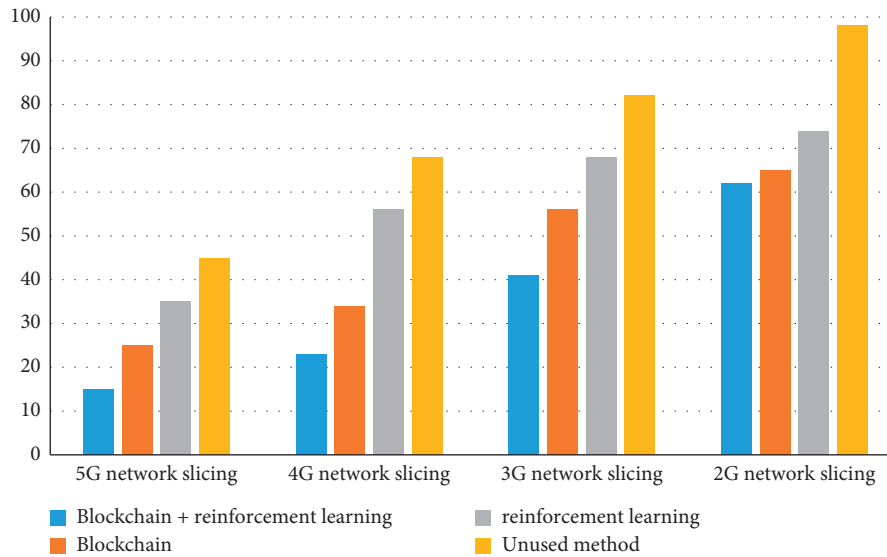


FIGURE 5: Delay comparison of different slice types.

The comparison results show that the delay increases with the increase of the number of slices, but the blockchain + reinforcement learning method has the lowest delay and can maintain the minimum delay. When the number of slices is 3, the delay is 155 ms. The overall delay of the blockchain is lower than the delay of reinforcement learning because the blockchain will give priority to nodes with rich resources and strong data processing capabilities when selecting nodes and link mappings, so the delay is lower.

4.1.2. Delay Comparison of Different Slice Types. Under different slice types, set the number of users to 30 and compare the delays generated by several methods. We compare 5G network slicing, 4G network slicing, 3G network slicing, and 2G network slicing in blockchain + reinforcement learning, blockchain, reinforcement learning, and unused methods. The results are shown in Figure 5.

Through the comparison results, it can be seen that the latency of 5G network slicing is lower than that of 4G, 3G, and 2G. 5G network slicing has the lowest latency of only 15 ms in the method of blockchain and reinforcement learning. This is because the greater the number of VNFs, the more nodes that the slice will pass through to process the same data packet, the longer the link that passes, and the greater the delay.

4.1.3. System Reliability. System reliability is an indispensable step before the experiment. We will compare the system reliability of different methods (blockchain + reinforcement learning, blockchain, reinforcement learning) under different numbers of users. The comparison result is shown in Figure 6:

It can be seen from the graph that the reliability decreases with the increase of the number of users because reliability is related to delay. The greater the transmission delay, the lower the reliability. The reliability of the

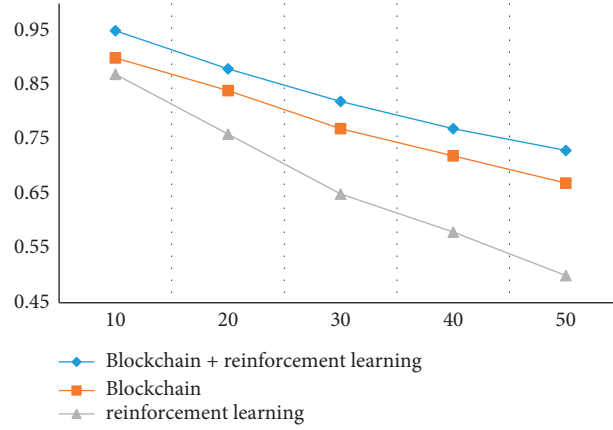


FIGURE 6: System reliability.

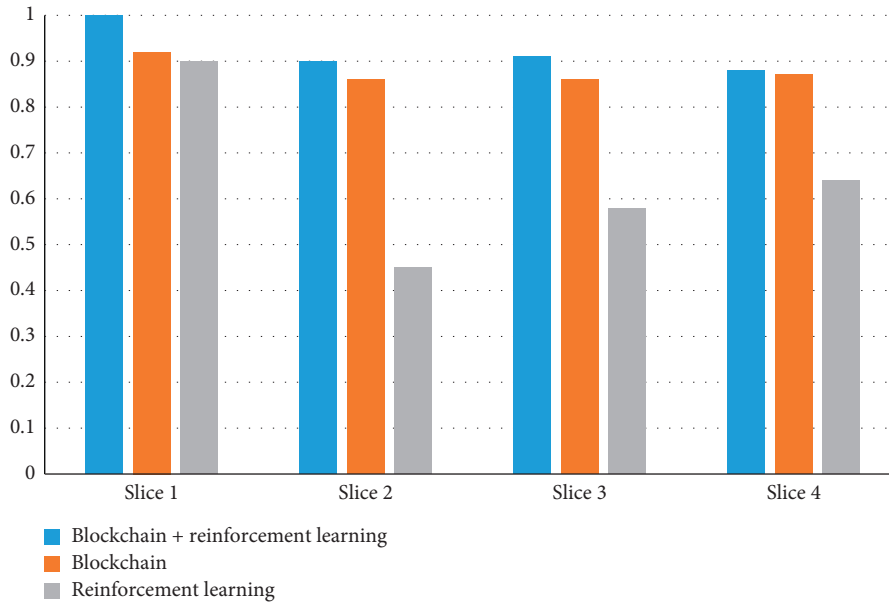


FIGURE 7: Resource utilization of different slices.

TABLE 5: Average cumulative receiving throughput.

Equipment	Blockchain + reinforcement learning	Blockchain	Reinforcement learning
Video stream 1	1450	1400	1150
Video stream 2	1350	1320	1160
IOT equipment	700	650	480
Mobile devices	270	260	100

supporting blockchain + reinforcement learning method is the highest, with a reliability of 0.95. This means that 5G network slicing that supports blockchain + reinforcement learning methods can provide services for more businesses.

4.2. Resource Utilization of Different Slices. This article studies the methods that support blockchain and reinforcement learning. We will study the resource utilization of blockchain and reinforcement learning for different

slices. Set up 4 slices and perform three tests on each slice, namely, blockchain + reinforcement learning, blockchain and reinforcement learning, and finally compare their resource utilization experiment results as shown in Figure 7:

According to the experimental results, it can be concluded that the method of blockchain + reinforcement learning has the highest resource utilization rate. The resource utilization rate of the four slices under the blockchain + reinforcement learning method is all above 0.8, and

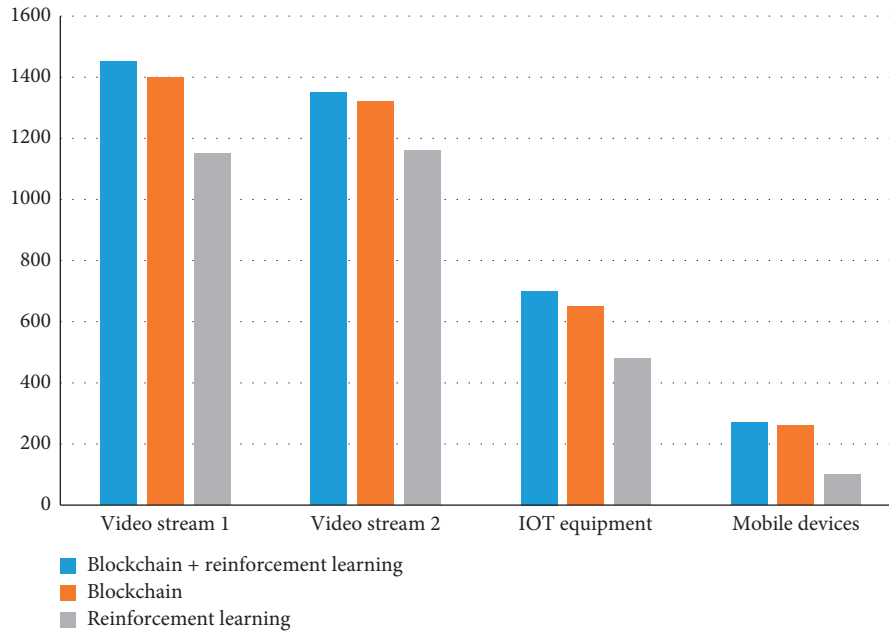


FIGURE 8: Average cumulative receiving throughput.

TABLE 6: Average QOE.

Equipment	Blockchain + reinforcement learning	Blockchain	Reinforcement learning
Video stream 1	0.83	0.75	0.68
Video stream 2	0.79	0.73	0.62
IOT equipment	0.7	0.64	0.47
Mobile devices	0.64	0.56	0.4

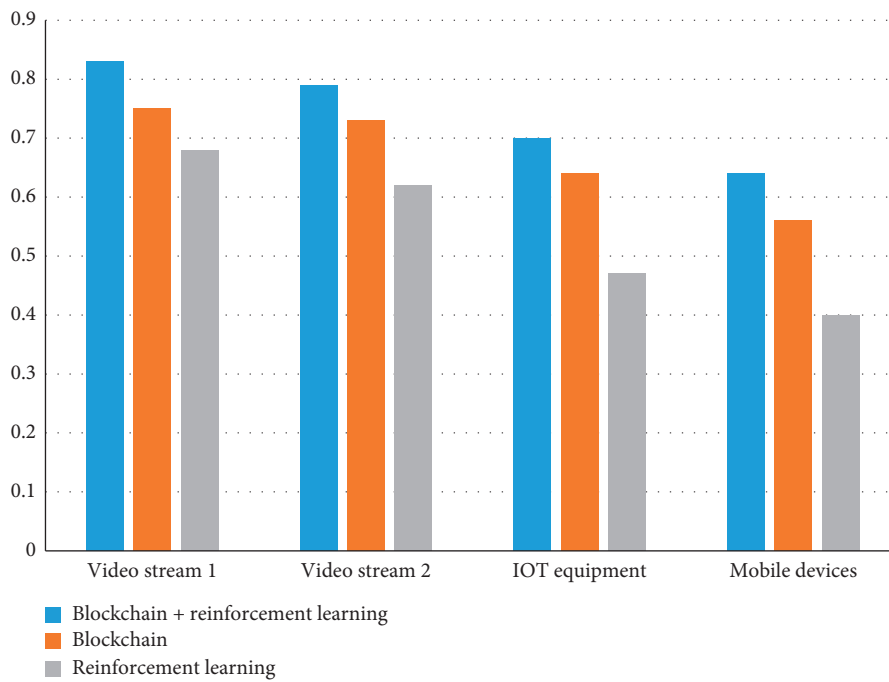


FIGURE 9: Average QOE.

the highest is 1. It shows that the block chain + reinforcement learning method has the best resource utilization for the block.

4.3. Simulation Test. According to the 5G network proposed in this article: Support blockchain and reinforcement learning methods to design a simulation test. Excluding other factors, the experimental subjects are video stream 1, video stream 2, IOT devices, and mobile devices.

4.3.1. Average Cumulative Receiving Throughput. The experiment will compare 4 types of equipment using three methods: blockchain + reinforcement learning, blockchain, and reinforcement learning. By comparing the average cumulative received throughput (kpbs), which method is better is decided. Throughput refers to the number of requests processed by the system in a unit of time. The results are shown in Table 5.

The result is plotted as a histogram, and the result is shown in Figure 8.

According to the experimental results, the average receiving throughput of video stream 1 is higher than that of video stream 2, IOT devices, and mobile devices, and the average cumulative receiving throughput is the highest under the blockchain + reinforcement learning method, reaching 1450 kpbs.

4.3.2. Average QOE. Under three different methods, compare the average QOE of different devices to prove which method is more suitable for 5G network slicing. QOE refers to the user's comprehensive experience of the quality and performance of the network system. The results are shown in Table 6:

The result is plotted as a histogram, and the result is shown in Figure 9.

According to the experimental results, the average QOE of video stream 1 is higher than that of video stream 2, IOT devices, and mobile devices, and the average QOE is the highest under the blockchain + reinforcement learning method, reaching 0.83.

5. Conclusion

With the advent of the 5G era, current technologies can no longer meet the needs of users. Network congestion and slow network speeds are major problems currently facing. In order for users to use network services more smoothly, network services are more convenient. This article designs 5G network slicing: a method model supporting blockchain and reinforcement learning. This model will perform better distribution management of the network, increase the transmission rate of users in the business, and reduce the transmission delay.

The research results of the article are given below:

- (1) In the model testing stage, the results of the study on the variation of the delay with the number of slices show that the delay increases with the increase of the number of slices, but the blockchain + reinforcement learning method has the lowest delay and can maintain the minimum delay. When the number of slices is 3, the delay is 155 ms.
- (2) The comparison of the delay of different slice types shows that the delay of 5G network slicing is lower than that of 4G, 3G, and 2G. 5G network slicing has the lowest delay in the method of blockchain and reinforcement learning, only 15 ms.
- (3) In the detection of system reliability, reliability decreases as the number of users increases. This is because reliability is related to delay. The greater the transmission delay, the lower the reliability. Supporting blockchain + reinforcement learning method has the highest reliability.
- (4) In the resource utilization experiment of different slices, it can be known that the method of blockchain + reinforcement learning has the highest resource utilization. The resource utilization rate of the four slices under the blockchain + reinforcement learning method is all above 0.8 and the highest is 1.
- (5) Through the simulation test of the experiment, the results show that the average receiving throughput of video stream 1 is higher than that of video stream 2, IOT devices, and mobile devices, and the average cumulative receiving throughput under the blockchain + reinforcement learning method is the highest, reaching 1450 kpbs. The average QOE of video stream 1 is higher than that of video stream 2, IOT devices, and mobile devices, and the average QOE is the highest under the blockchain + reinforcement learning method, reaching 0.83.

Although the results of this experiment are obvious, it has certain limitations and is limited to the use of 5G network slicing. A lot of research is needed in the future to enhance its universality and apply it to more scenarios. In future research, the methods for supporting blockchain and reinforcement learning proposed in this article can be improved, so that blockchain and reinforcement learning methods can be realized in the network service requirements with more goals.

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 regarding this work.

References

- [1] H. Zhang, N. Liu, X. Chu, K. Long, A.-H. Aghvami, and V. C. M. Leung, "Network slicing based 5G and future mobile networks: Mobility, resource management, and challenges," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 138–145, 2017.
- [2] J. Ordóñez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca, and J. Folgueira, "Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80–87, 2017.
- [3] M. Jiang, M. Condoluci, and T. Mahmoodi, "Network slicing management & prioritization in 5G mobile systems," in *Proceedings of the European Wireless 2016; 22th European Wireless Conference*, VDE, Oulu, Finland, May 2016.
- [4] M. Luo, J. Wu, and X. Li, "Cross-domain certificateless authenticated group key agreement protocol for 5G network slicings," *Telecommunication Systems: Modelling, Analysis, Design and Management*, vol. 74, 2020.
- [5] M. A. Kourtis, T. Sarlas, and G. Xilouris, "Conceptual evaluation of a 5G network slicing technique for emergency communications and preliminary estimate of energy trade-off," *Energies*, vol. 14, 2021.
- [6] C. Campolo, A. Molinaro, and A. Iera, "5G network slicing for vehicle-to-everything services," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 38–45, 2018.
- [7] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, K. Samdanis, and X. Costa-Perez, "Optimising 5G infrastructure markets: the business of network slicing," in *Proceedings of the IEEE Infocom-IEEE Conference on Computer Communications*, pp. 1–9, IEEE, Atlanta, GA, USA, May 2017.
- [8] M. Jiang, M. Condoluci, and T. Mahmoodi, "Network slicing in 5G: An auction-based model," in *Proceedings of the 2017 IEEE International Conference on Communications (ICC)*, IEEE, Paris, France, May 2017.
- [9] R. Pries, H. J. Morper, N. Galambosi, and M. Jarschel, "Network as a service-a demo on 5G network slicing," in *Proceedings of the Teletraffic Congress*, IEEE, Germany, September 2017.
- [10] Y. Dai, D. Xu, S. Maharjan, Z. Chen, Q. He, and Y. Zhang, "Blockchain and deep reinforcement learning empowered intelligent 5G beyond," *IEEE Network*, vol. 33, no. 3, pp. 10–17, 2019.
- [11] N. N. Ahamed and P. Karthikeyan, "A reinforcement learning integrated in heuristic search method for self-driving vehicle using blockchain in supply chain management," *International Journal of Intelligent Networks*, vol. 1, 2020.
- [12] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "Privacy-preserved task offloading in mobile blockchain with deep reinforcement learning," *IEEE Transactions on Network and Service Management*, vol. 99, p. 1, 2020.
- [13] L. Long-Ji, "Self-improving reactive agents based on reinforcement learning, planning and teaching," *Machine Learning*, vol. 8, 1992.
- [14] A. Kalokylos, "A survey and an analysis of network slicing in 5G networks," *IEEE Communications Standards Magazine*, vol. 2, no. 1, pp. 60–65, 2018.
- [15] Q. Jia, R. Xie, T. Huang, J. Liu, and Y. Liu, "Efficient caching resource allocation for network slicing in 5G core network," *IET Communications*, vol. 11, no. 18, pp. 2792–2799, 2017.
- [16] P. Stone, R. S. Sutton, and G. Kuhlmann, "Reinforcement learning for RoboCup soccer keepaway," *Adaptive Behavior*, vol. 13, no. 3, pp. 165–188, 2005.
- [17] G. Chen, J. Qi, C. Tang, Y. Wang, Y. Wu, and X. Shi, "Analysis and research of key genes in gene expression network based on complex network," *Complexity*, vol. 2020, Article ID 8816633, 12 pages, 2020.
- [18] 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, 2021.
- [19] A. Theodorou, J. Buchli, and A. L. S. Sch, "A generalized path integral control approach to reinforcement learning," *Journal of Machine Learning Research*, vol. 11, no. 11, pp. 3137–3181, 2010.
- [20] Y. Yuan and F. Y. Wang, "Blockchain: The state of the art and future trends," *Acta Automatica Sinica*, vol. 42, 2016.
- [21] L. Sun, Q. Yu, D. Peng, S. Subramani, and X. Wang, "FogMed: a fog-based framework for disease prognosis based medical sensor data streams," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 603–619, 2021.
- [22] K. Verbert, M. Sharples, and T. Klobuar, "The blockchain and kudos: A distributed system for educational record, reputation and reward," in *Proceedings of the European Conference on Technology Enhanced Learning*, Springer International Publishing, Lyon, France, 2016.
- [23] M. H. Abidi, H. Alkhalefah, and K. Moiduddin, "Optimal 5G network slicing using machine learning and deep learning concepts," *Computer Standards & Interfaces*, vol. 76, no. 1, Article ID 103518, 2021.
- [24] C. H. Tobar, A. Ordóñez, and O. Rendon, "Scalability and performance analysis in 5G core network slicing," *IEEE Access*, vol. 8, pp. 142086–142100, 2020.
- [25] L. Tang, Y. Zhou, and Y. Yang, "Virtual network function dynamic deployment algorithm based on prediction for 5G network slicing," *Journal of Electronics and Information Technology*, vol. 41, no. 9, pp. 2071–2078, 2019.

Research Article

An Empirical Analysis of the Influence of Volleyball Elective Course on Students' Physical Health Based on Digital Image

Shiwei Wang 

Chengdu Sport University, Chengdu, Sichuan 610041, China

Correspondence should be addressed to Shiwei Wang; 100373@cdsu.edu.cn

Received 26 January 2022; Revised 17 February 2022; Accepted 21 February 2022; Published 24 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Shiwei Wang. 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 recent years, due to the continuous improvement of the national economic level and the increasing academic burden of students' main courses, students' physical health problems (e.g., obesity, vision, and lumbar spine) have become more and more serious, which urgently needs the attention of relevant departments of national education and parents. This paper will use digital image technology to create a physical parameter measurement system and use literature, comparative analysis, and other research methods to analyze the impact of volleyball elective courses on students' physical health. Firstly, this paper explains the theory of image processing technology and analyzes the parameters of human body scientifically; secondly, it detects the physical parameters of human body in digital images and also designs an image recognition system; finally, through experimental analysis, the accuracy of identifying key points of images is relatively high. After the system is adopted, the error of the measurement index is small. After the training of human body indexes, the effect of volleyball can be effectively improved.

1. Introduction

Students are the future pillars of a country and the hope and seeds of a country. However, with the rapid development of science and technology, students enjoy the convenience of life by using mobile phones, computers, and other electronic devices for a long time. Long-term inactivity (or a small amount of exercise) leads to students' insufficient physical fitness and gradual decline in physical fitness, which is contrary to our concept of "sunshine sports." In order to cultivate students' idea of "lifelong physical education and physical fitness," this paper selects volleyball elective course to test the specific role of students' health, using digital image technology to analyze and study, discuss and help students to make a better fitness plan suitable for individuals from various angles and factors, truly make people-oriented, provide effective scientific basis and health guidance suggestions, and strive to improve the physique of all students. Reference [1] introduces the basis of digital image, image enhancement in spatial domain and frequency domain. Literature [2] uses a class of algorithms to determine the local similarity between structured data sets as the basis of

pattern recognition and image processing. Literature [3] investigates and analyzes digital image recording methods and reviews common standards and guidelines. Reference [4] combines artificial neural network with digital image to analyze grain quality. Reference [5] explains the use of digital image related theory and algorithm structure in processing algorithms and applications. Literature [6] used experiments to find out the comprehensive effects of physical fitness and psychological variables on abdominal strength, grip and psychological skills of volleyball players after training. Literature [7] showed the training effect of explosive leg force and shoulder strength of volleyball players after training by comparing the situation after training in Kupas. Literature [8] shows that the application of stair teaching method in volleyball elective courses in colleges and universities can effectively improve students' comprehensive volleyball ability. Literature [9] uses algorithms to detect the motion features based on body trajectory clustering and aggregation in volleyball and collects the required motion data. Literature [10] introduces the effective improvement of intensity modified leisure volleyball on the health signs and physical fitness of men aged 25–55. Literature [11] discusses the

influence of four-person soft volleyball teaching practice, which is beneficial to college students' physical and mental health. Literature [12] reviews and analyzes the research on students' mastery of health-related fitness knowledge so as to clear up misunderstandings. Literature [13] investigated the attitudes and physical activity behaviors of professionals before participating in sports, health, and leisure services. Literature [14] evaluates the body composition, health awareness, and cardiopulmonary health of female college students after compulsory physical education class. Literature [15] studies the theoretical and empirical analysis of the influence of age and exercise on physical activity. The article introduces the image processing technology to carry on the movement recognition to the volleyball movement; the human body physique characteristic and the scientific movement training through the high recognition rate can be scientifically analyzed. Analysis of the human body constitution in the article shows that, after training, the body shape is more increased than before, and the overall weight shows a downward trend. Volleyball elective course has different degrees of improvement on each student's physical fitness, and to achieve the best results, so as to strengthen the diversified teaching practice.

2. Theoretical Basis

2.1. Research Methodology. This paper mainly adopts the following three methods to study, analyze, and compare the theory and data in order to ensure the reliability, authenticity, and effectiveness of the results and avoid adverse effects caused by other unexpected factors.

2.1.1. Literature Research Method [16]. This paper refers to a large number of websites and documents such as CNKI, China Academic Journal Network, and Baidu Academic, which provides scientific theoretical basis for the study of students' physical health. In addition, this paper also consulted a large number of books, periodicals, lectures, conferences, and magazines related to physical education quality courses, physical education teaching quality reform, volleyball elective courses, and so on, in order to deeply understand and grasp the research content of this topic from all aspects and angles.

2.1.2. Data Statistics [17]. As the name implies, this method is to collect and sort out all kinds of data obtained from experimental tests so as to facilitate the analysis and discussion of the results. This method will use Excel 2016 software [18] and SPSS 17.0 tool [19] to collect various indicators for statistical steps.

2.1.3. Comparative Analysis [20]. According to different categories and requirements, the physical health indicators of students obtained after the experimental test are compared and analyzed, such as body shape indicators and physical fitness indicators. Finally, the conclusion is summarized.

2.2. Digital Images

2.2.1. Introduction of Concepts. A digital image [21] is one of the images, represented by a two-dimensional array [22], and the basic digital unit is pixel points [23], which can be directly processed by a computer. The other analog image cannot be processed by computer unless it is digitally converted. The image format used in this paper is mainly JPEG format, which is mainstream and can save space by compressing storage, and basically all shooting equipment can support it. In this paper, according to the brand effect and use effect, we finally decide to use a certain Sony device to complete the shooting work. As shown in Figure 1, it is about the representation coordinate system of pixels on the image.

2.2.2. Instructions for Image Acquisition. In this paper, firstly, the two-dimensional plane image is taken to ensure the high quality of the image quality, and the image parameter standard is unified, and the two-dimensional plane image is converted into the parameter data of the three-dimensional index.

2.2.3. Image Processing Techniques. (1) *Image Grayscale [24].* It is a general popular processing technique that relies on RGB [25] color proportions to operate. It can improve the speed of image processing, transform the original color image with too large pixels into gray image with fewer pixels, and reduce the burden of image processing. As shown in Figures 2 and 3, it is a spatial representation of the relationship between gray value number axis and RGB color scale.

As shown below, there are several conversion methods related to image grayscale. According to the actual needs and convenience, we chose the fifth method to effectively separate human parameters from surrounding environmental parameters.

Method 1: floating-point algorithm.

$$Gray = R * 0.3 + G * 0.59 + B * 0.11. \quad (1)$$

Method 2: integer method.

$$Gray = (R * 30 + G * 0.59 + B * 11) / 100. \quad (2)$$

Method 3: shift algorithm.

$$Gray = (R * 76 + G * 151 + B * 28) \gg 8. \quad (3)$$

Method 4: average method.

$$Gray = (R + G + B) / 3. \quad (4)$$

Method 5: only take the green method.

$$Gray = G. \quad (5)$$

Suppose a grayscale photo of $M \times N$ size; list all possible grayscale values $z_i, i = 0, 1, 2, \dots, L - 1$. We can get the probability of gray level z_k in this photo:

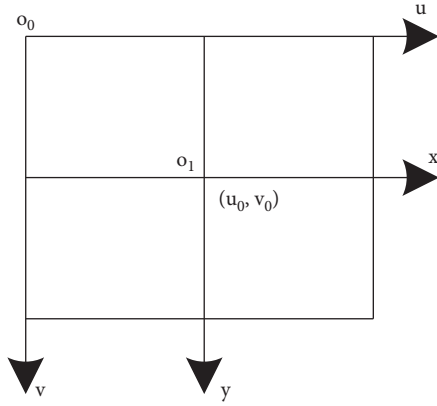


FIGURE 1: The coordinates of the image.

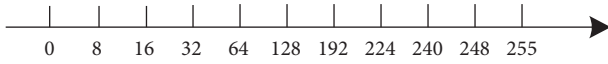


FIGURE 2: Gray value.

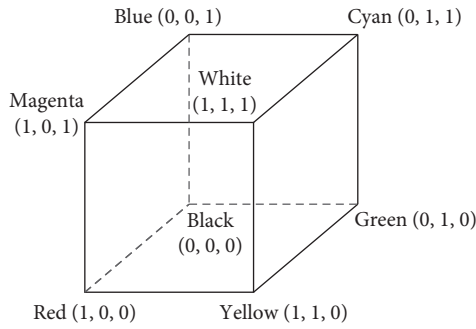


FIGURE 3: RGB color space.

$$p(z_k) = \frac{n_k}{MN}, \quad (6)$$

$$\sum_{k=0}^{L-1} p(z_k) = 1. \quad (7)$$

From formulas (6) and (7), the average grayscale, grayscale variance, and n -step distance can be obtained:

$$m = \sum_{k=0}^{L-1} z_k p(z_k), \quad (8)$$

$$\sigma^2 = \sum_{k=0}^{L-1} (z_k - m)^2 p(z_k), \quad (9)$$

$$\mu_n(z) = \sum_{k=0}^{L-1} (z_k - m)^n p(z_k). \quad (10)$$

(2) *Image Arithmetic Operation.* If the image is disturbed by Gaussian noise, superposition operation is performed:

$$s(x, y) = f(x, y) + g(x, y). \quad (11)$$

Image enhancement operations are

$$d(x, y) = f(x, y) - g(x, y). \quad (12)$$

Select a special image block:

$$p(x, y) = f(x, y) \times g(x, y). \quad (13)$$

Because the shot lens is affected by optical characteristics, shadows will occur. The lens principle is as shown in Figure 4.

We perform shadow correction of the image:

$$v(x, y) = \frac{f(x, y)}{g(x, y)}. \quad (14)$$

(3) *Geometric Space Transformation.* It can improve the spatial relationship of coordinates between pixels, where (v, w) = original pixel coordinates and (x, y) = transformed pixel coordinates.

$$(x, y) = T\{(v, w)\}. \quad (15)$$

(4) *Image Transformation.* The goal is to transform into a more convenient form by calculation, and then transform back, which can effectively improve the processing efficiency.

Unified formula of transformation is

$$T(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) r(x, y, u, v). \quad (16)$$

Corresponding inverse transformation unified formula is

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u, v) s(x, y, u, v). \quad (17)$$

For example, the process of filtering to remove interference sources is illustrated:

$$r(x, y, u, v) = e^{-j2\pi(ux/M+vy/N)}, \quad (18)$$

$$s(x, y, u, v) = \frac{1}{MN} e^{j2\pi(ux/M+vy/N)}. \quad (19)$$

Formulas (18) and (19) are substituted into formulas (16) and (17), respectively, and the result of transformation is as follows:

$$T(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M+vy/N)}, \quad (20)$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u, v) e^{j2\pi(ux/M+vy/N)}. \quad (21)$$

(5) *Image Edge Detection.* As shown in Figure 5, it is an introduction to different types of edge shapes.

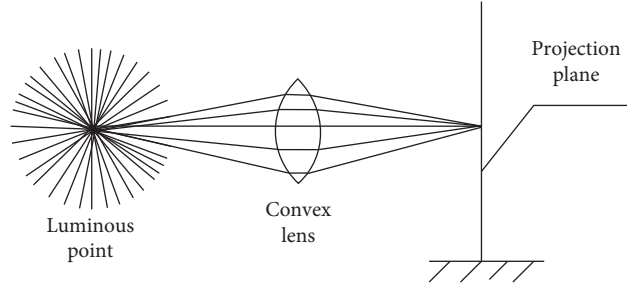


FIGURE 4: Principle of lens shadow generation.



FIGURE 5: Edge of gray image. (a) Step-like. (b) Pulsatile.

The principle of image edge detection is to identify the change of gray value in image edge. Based on Laplacian operator, its template is as Figure 6.

The relevant formula is as follows:

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}, \quad (22)$$

$$\nabla^2 f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y). \quad (23)$$

(6) *Image Binarization*. If we want to detect our target object from the initial image, we use this method.

$$g(x, y) = \begin{cases} 0 & (\text{Gray value is less than threshold } (T)) \\ 255 & (\text{Gray value is greater than threshold } (T)) \end{cases} \quad (24)$$

In formula (24), *threshold* (T) is the judgment threshold, and the corresponding optimal value is generally obtained by continuously comparing performance parameters in experimental comparison.

2.3. Parameter Indicators

2.3.1. *Confirmation of Human Body Size Proportion*. Using digital image technology to measure the human body, the obtained data is calculated in equal proportion with the defined human body size ratio, and the original real human body data is restored. The defined size ratio is shown in Table 1.

The formula model for obtaining data from the formula of human cross section model is as follows:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

FIGURE 6: Laplacian operator template.

TABLE 1: Size ratio of each part of the human body.

Position	Boy	Girl
Height	1.00	1.00
Eyes	0.94	0.93
Neck	0.84	0.86
Shoulder	0.82	0.81
Axillary	0.75	0.75
Chest circumference	0.72	0.72
Diaphragm	0.70	0.70
Elbow	0.63	0.63
Waist circumference	0.61	0.63
Hip circumference	0.53	0.53
Crotch	0.47	0.47
Thigh	0.41	0.41
Knee	0.26	0.26

$$AX^{2n} + BX^n Y^n + CY^n = Z^{2n}, \quad (25)$$

$$X = (R + S_x \times \sin \theta) \times \cos \theta,$$

$$Y = R \times \sin \theta \times \alpha. \quad (26)$$

2.3.2. *Unit Conversion*. Because the unit of image is pixel unit, to convert it into actual human body parameters, it is necessary to convert pixel unit into actual centimeter unit.

This conversion is based on 100 cm as a unit length, and the conversion formula is as follows:

$$L_{(Actual)} = L_{(Pixel)} * \left(100/L_{(Ruler)}\right)cm. \quad (27)$$

In the second part of the article, the related research methods are explained, and then the digital image concept, acquisition, and processing technology are explained, and then the human body parameter characteristic parameters are explained proportionally. The second part gives an effective explanation of the basic research work of the article, which has good scientific significance.

3. Physical Parameter Detection Based on Digital Images

3.1. Analysis of the Characteristics of Volleyball Elective Courses. Volleyball elective course has antagonism across the net, competition and skill, which can stimulate students to learn volleyball elective course more passionately and enhance their physique. Volleyball class requires students to be skilled in volleyball basic skills, solid and stable core strength, know the rules of sports venues and referee rules of volleyball matches, skillfully use their own physical strength and skills, and be able to react quickly to volleyball, and meet the requirements of eyesight and self-control. Therefore, in volleyball elective training, physical education teachers only think about how to enhance students' physical fitness reasonably and effectively in order to meet the requirements of students' activities inside and outside class. As shown in Table 2, a key explanation about volleyball elective course is presented.

3.2. Analysis of Students' Physique. In recent years, with the rapid development of the world, the rising power of science and technology has brought many changes in life. These changes make people's lives more convenient, but also bring many troubles to people, among which physical health problems are particularly prominent. Due to the long-term academic pressure, students' spare time is occupied by homework and electronic equipment, and only a few students take physical exercise, which leads to the gradual decline of students' physical health and presents a subhealth state, which deserves the attention of relevant state departments and parents, and time should be given to encourage extracurricular activities.

When we test, we will use various indicators (e.g., height, weight, and vital capacity) to reflect the basic health status of a student according to the national test standards and make reasonable analysis and suggestions.

3.3. Design and Implementation of the System

3.3.1. Functional Design of the System. The system designed in this paper is a tool for measuring human body in experimental test, adopting advanced digital image technology, using shooting equipment to record the image of human body, and using related image processing technology to preprocess the image. It includes a series of operations such

TABLE 2: Characteristics of volleyball elective courses.

Name	Volleyball
Basic technology	Step, double pad, serve
Specialized technology	Application of tactical skills
Physical quality	Speed, strength, agility

as image grayscale, image arithmetic operation, image binarization, image transformation, image edge detection, and so on, and truly restores and converts the real data of human body as the data reference of the experiment, which will liberate the tedious and repeated work in the experiment and automatically obtain human body parameters. A flowchart of the image processing of the system is shown in Figure 7.

3.3.2. Implementation of the System

- (1) LabVIEW 2013 software in Windows 10 operating system is selected as the development language. This programming language is more concise than the traditional programming language; there are more practical functional modules, more suitable for this system design needs.
- (2) After the size extraction, the system will use Excel 2016 software and SPSS 17.0 tools to collect various indicators for statistical steps so that the data can be saved more conveniently and reflected intuitively, which is convenient for us to analyze.
- (3) Realize the design drawing.

First, temporarily create a memory location to store the images we need to wait for processing. The functional design of the image reading part is shown in Figure 8.

The types of images are shown in Table 3.

Second, the purpose of this program is to obtain the B primary color (color image), and the functional design diagram of extracting grayscale is shown in Figure 9.

Among them, the pattern of color is shown in Table 4.

Third, let the value of B primary color become the gray value of each pixel; we can use MATLAB code to achieve gray image. As shown in Algorithm 1, it is about the process realized by MATLAB code, and we choose three methods to run it.

Fourthly, the acquired gray image is processed, and the binarization of the image is carried out by MATLAB, as shown in Figure 10.

Fifthly, all kinds of data are uniformly converted into strings and then stored in corresponding files. The image data saving function design diagram is shown in Figure 11.

4. Experimental Data Analysis

4.1. Subjects of the Study. This paper takes the influence of volleyball elective course on students' physical health as the research object and makes a comparative study on the changes of students' physical health before and after volleyball elective course. Because the student group includes kindergarten, primary school, junior high school, high school, and university, in order to avoid the influence of

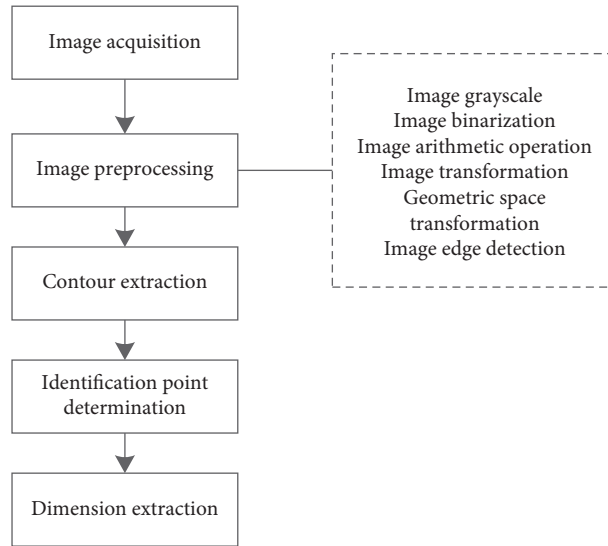


FIGURE 7: System operation flowchart.

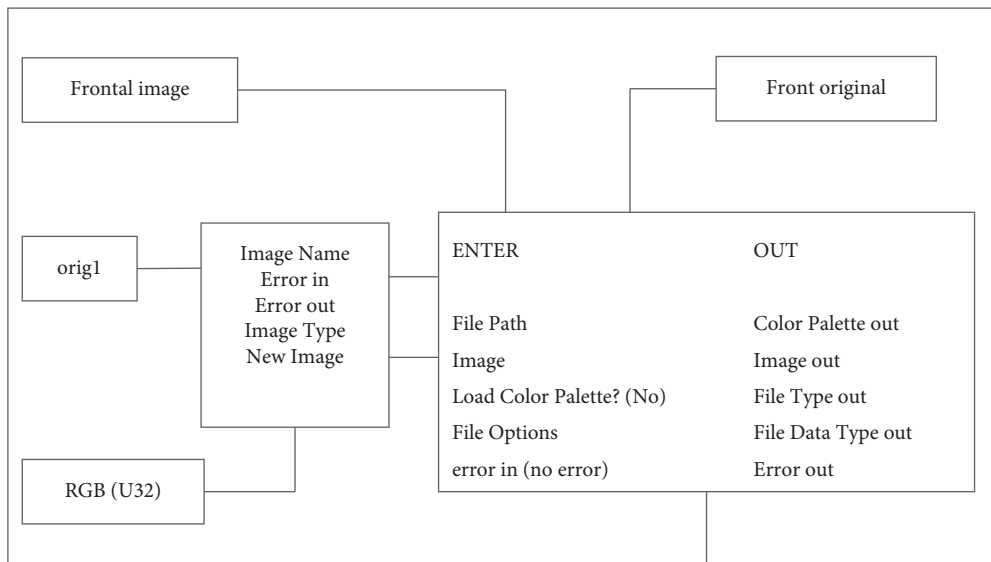


FIGURE 8: Design of reading function.

TABLE 3: Table of image types.

Type	Size per pixel
Gray scale (U8)	8 bits (unsigned, standard monochrome)
Gray scale (16)	16 bits (signed)
Gray scale (SGL)	32 bits (floating point type)
Gray scale (U16)	16 bits (unsigned, standard monochrome)
Composite (CSG)	2 * 32 bits (floating-point type)
RGB (U32)	32 bits (red, green, blue, transparency)
HSL (U32)	32 bits (chromaticity, saturation, brightness, transparency)
RGB (U64)	64 bits (red, green, blue, transparency)

physical development on students' health, we choose full-time freshmen in a university as the research object and randomly select 100 students (half male and half female) to

participate in the physical health test. The test is divided into three stages. The first stage is that all people have not participated in volleyball elective courses for the first time,

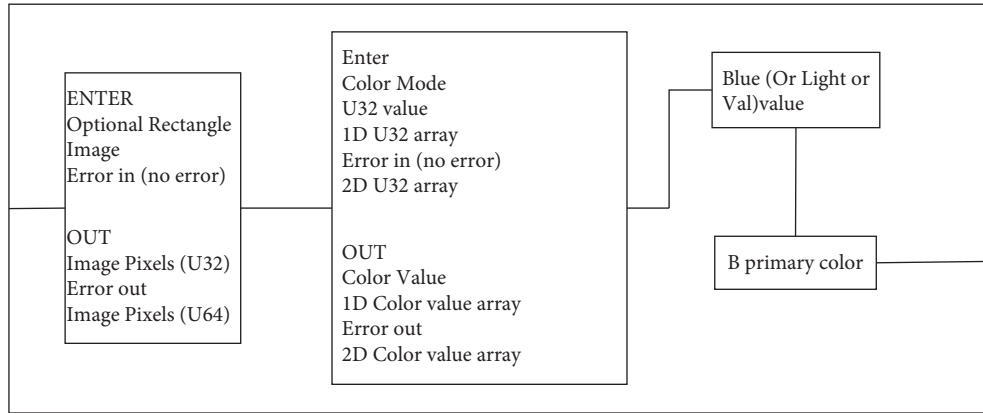


FIGURE 9: Design of gray level extraction function.

TABLE 4: Color pattern table.

RGB	Default value
HSL	Conversion to HSV
HSV	Conversion to HSV
HSI	Conversion to HSV

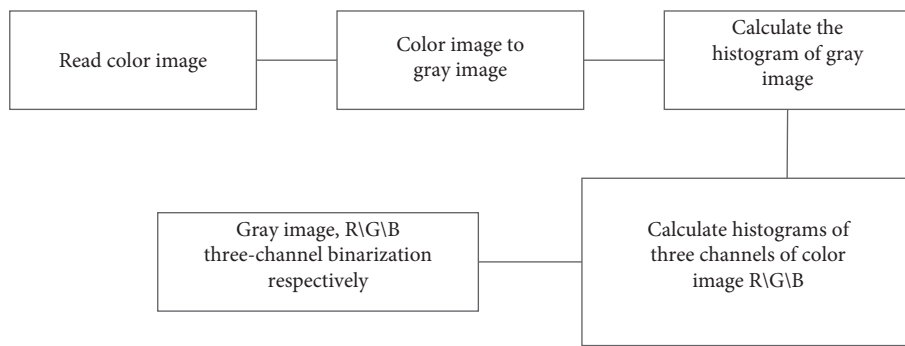


FIGURE 10: Image binarization flow.

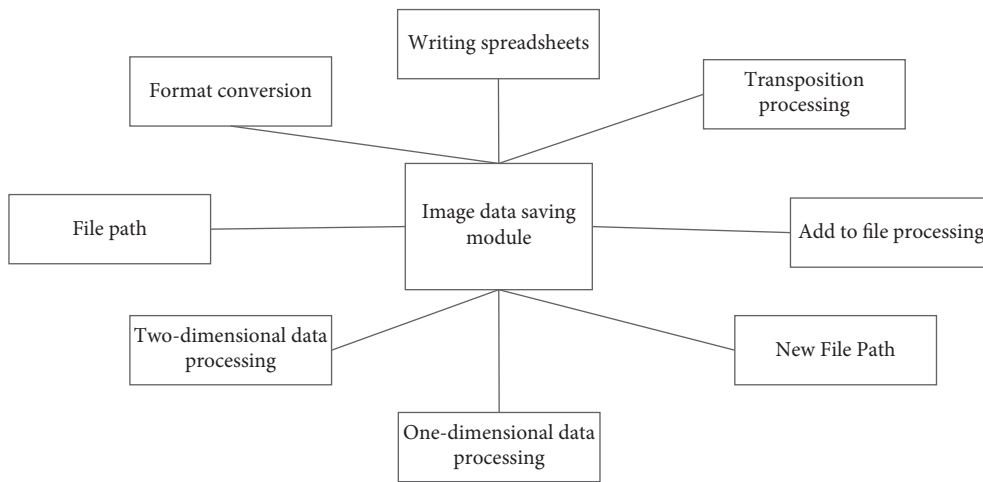


FIGURE 11: Design of image data saving function.

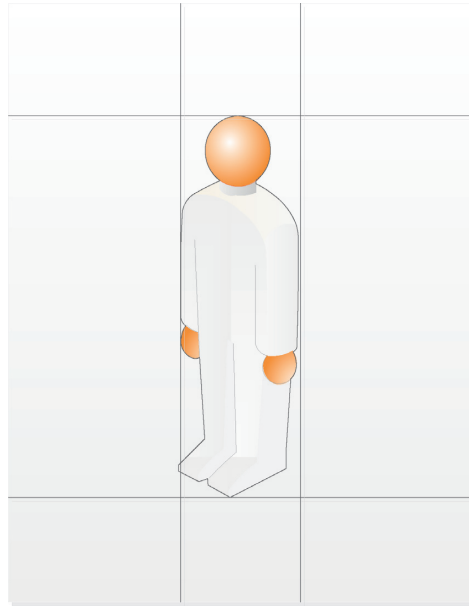


FIGURE 12: Calibration drawing of human body.

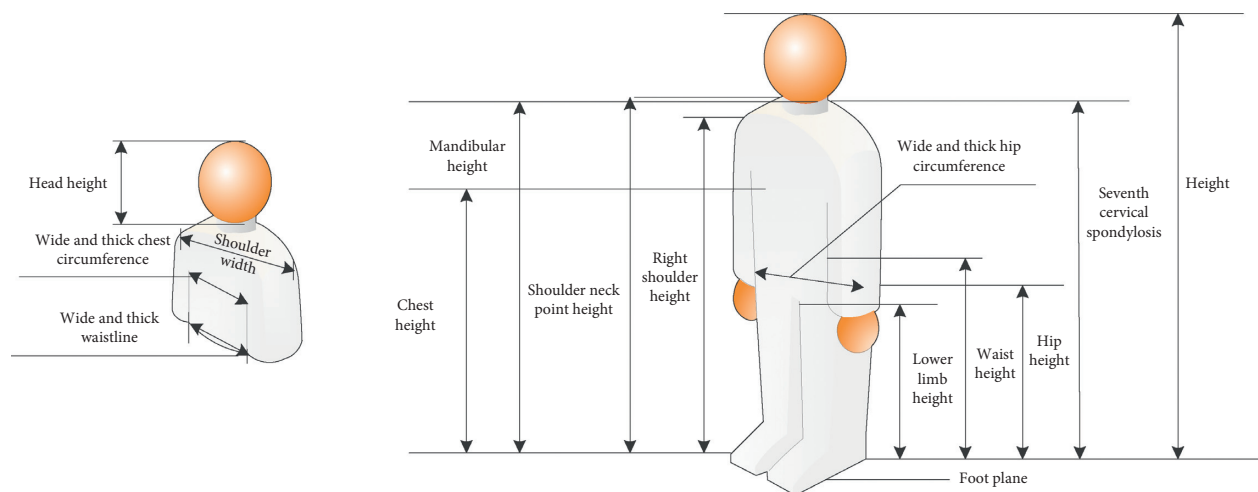


FIGURE 13: Key points of human body scanning.

the second stage is that all people have been tested for the second time after two years of volleyball exercise with the same intensity (they can no longer participate in other physical exercises), and the third stage is the third test after stopping training for one year. Due to space constraints, we only choose to show part of the test data.

4.2. Analysis of Image Processing

4.2.1. Image Scanning Determination

- (1) An example of a scan line determination diagram of a human body is shown in Figure 12.
- (2) The confirmation of key points in human body scanning is shown in Figure 13.

4.2.2. Scanning Recognition Accuracy Analysis. When processing digital images, it is very important to accurately confirm key points. Therefore, we analyzed the data of whether the digital image can accurately identify the key points in the human body and conducted five tests, respectively, as shown in Figure 14.

We calculate the average of the test results of five times and show the results in the form of line chart, as shown in Figure 15.

Analysis. We can find that the accuracy of digital image recognition is over 94%; the highest can reach 100%. The experimental results show that the accuracy of key points in image recognition is very high, which can meet the needs of the experiment.

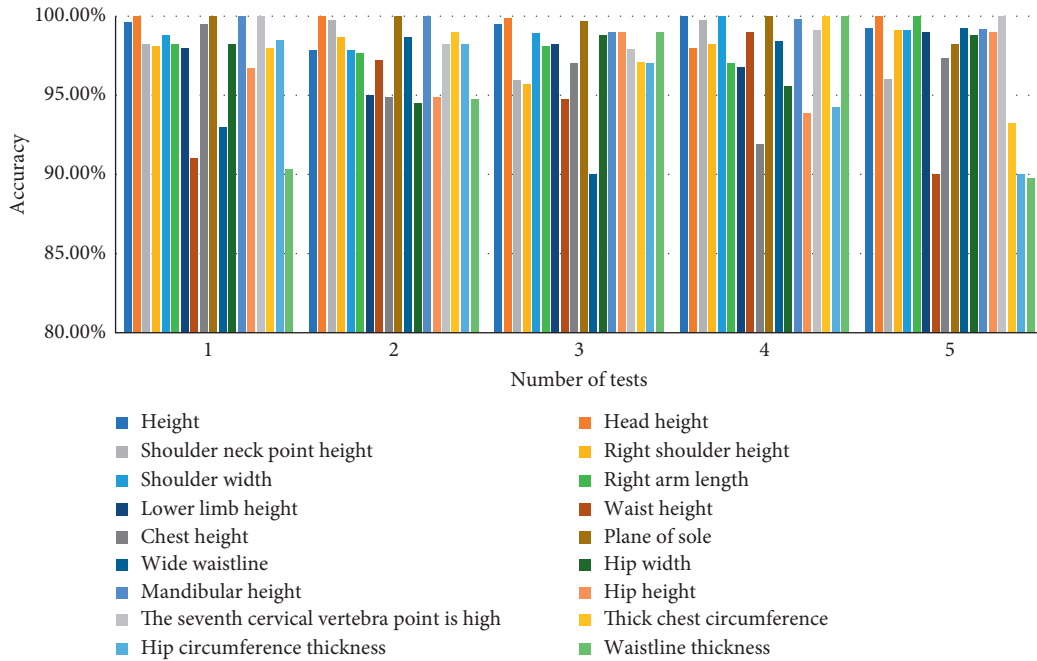


FIGURE 14: Recognition accuracy.

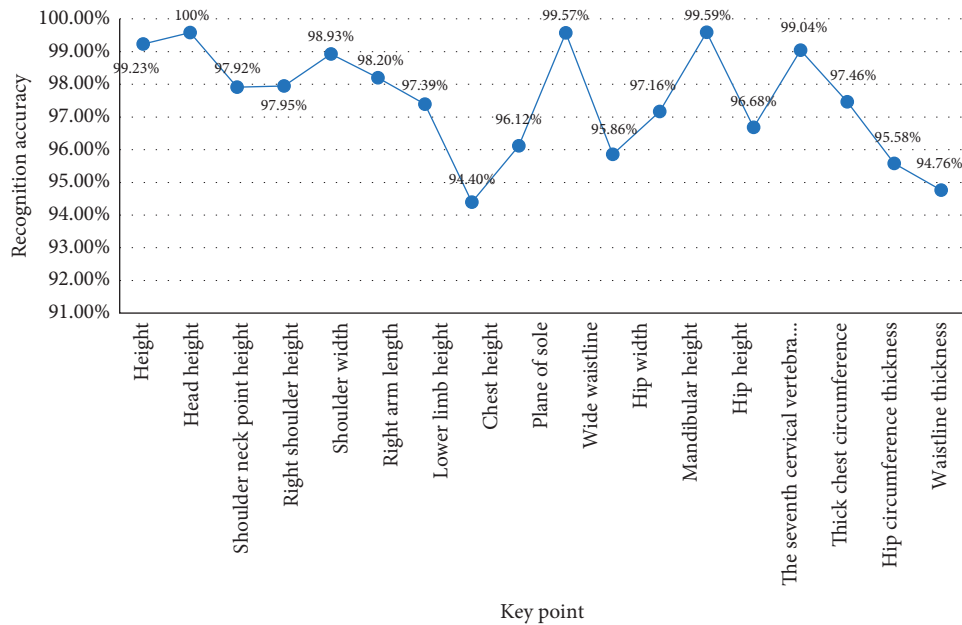


FIGURE 15: Average accuracy.

4.3. *Data Reliability Analysis.* The data detected by our system still need to be verified and the reliability of the data is analyzed. Only by comparing the error between the data tested by the system and the data measured by the real system can we judge that the measured data of our system is true, reliable, and effective. Otherwise, the system cannot be adopted and needs to be redesigned and improved. In this trial run of the system, we considered labor cost, time cost, and other reasons, and only selected

ten students for this data test, and adopted anonymous methods to protect students' privacy. However, in order to ensure the accuracy of the manual test results, we promise to repeat the test three times to select the average value.

As shown in Table 5, it is about the comparison of male students' height and arm span length.

As shown in Table 6, it is about the comparison of female students' height and arm span length.

```

Input: Img=imread('*.jpg');
Output: imshow()
[n m a]=size(Img);
GrayImage= rgb2gray(Img);
Img_Gray=zeros(n, m);
for x=1:n
    for y=1:m
        Img=max(Img(x, y, 1), max(Img(x, y, 2), Img(x, y, 3)));
        GrayImage =(Img (x, y, 1)+Img (x, y, 2)+Img (x, y, 3))/3;
        Img_Gray=0.3 * Img (x, y, 1)+0.59 * Img (x, y, 2)+0.11 * Img (x, y, 3);
    end
end
figure.imshow(Img);
figure.imshow(GrayImage);
figure.imshow(uint8(Img_Gray));

```

ALGORITHM 1: Gray image by MATLAB code.

TABLE 5: Data table of male students' height and arm span (unit: cm).

Sample number	Height (automatic)	Height (salary)	Arm extension (automatic)	Arm extension (artificial)
1	173.77	173.56	167.52	167.36
2	180.01	179.98	175.85	175.45
3	175.33	175.02	175.85	175.24
4	176.37	176.01	172.21	171.98
5	169.03	168.96	165.44	164.95
6	174.81	174.25	168.04	167.36
7	187.30	187.02	172.73	172.56
8	167.52	166.95	149.30	149.40
9	174.81	174.25	163.88	163.75
10	173.77	173.54	167.52	167.39

TABLE 6: Data table of height and arm span of girls (unit: cm).

Sample number	Height (automatic)	Height (salary)	Arm extension (automatic)	Arm extension (artificial)
1	160.23	159.98	158.67	157.39
2	161.25	160.89	160.55	159.68
3	159.45	159.35	159.33	159.54
4	152.27	152.45	152.48	152.98
5	153.68	153.65	153.83	153.22
6	164.72	164.54	164.17	163.95
7	150.73	150.05	150.26	150.95
8	169.59	169.32	168.65	168.06
9	157.49	157.24	157.30	157.65
10	158.23	158.24	158.67	158.66

As shown in Table 7, it is a comparison of the data of male students' measurements (chest circumference, waist circumference, and hip circumference).

As shown in Table 8, it is a comparison of the measurements (chest circumference, waist circumference, and hip circumference) of female students.

Conclusion. Through the comparison of the above table data, we can find that the error of each measurement index is very small, basically within the range of 0.2 cm, which can meet the needs of students' physical health detection in our experiment. After the system is adopted, we can find that its efficiency will be far greater than that of manual measurement.

4.4. Comparative Analysis of Data. Because the data of 100 students was tested, the data was too huge, so in the comparative analysis and display of the data here, we only randomly intercepted a small number of students' data (which was inconsistent with the student sample number when testing the system, and renumbered it) and adopted anonymous methods to protect students' privacy.

4.4.1. Comparative Analysis of Basic Physical Indicators. (1) We compare and analyze the data of height and arm span of male students and female students, respectively, and give the corresponding line chart, which is convenient to feel the

TABLE 7: Data sheet of male students' measurements (unit: cm).

Sample number	Bra (automatic)	Bra (artificial)	Waist circumference (automatic)	Waist circumference (salary)	Hip circumference (automatic)	Hip circumference (artificial)
1	83.41	83.44	81.99	82.01	88.63	88.94
2	106.22	105.84	83.22	83.71	95.52	95.47
3	84.16	84.28	77.32	77.53	79.91	80.11
4	88.66	88.49	79.21	79.33	85.54	85.31
5	80.13	80.09	75.88	75.69	77.92	77.87
6	97.40	97.25	78.36	78.17	85.67	85.58
7	113.11	112.89	84.30	84.31	108.15	108.04
8	91.31	91.48	84.78	84.59	99.31	99.21
9	81.09	80.83	71.21	71.06	87.88	87.64
10	108.34	108.29	75.35	75.24	111.13	110.97

TABLE 8: Data sheet of girls' measurements (unit: cm).

Sample number	Bra (automatic)	Bra (artificial)	Waist circumference (automatic)	Waist circumference (salary)	Hip circumference (automatic)	Hip circumference (artificial)
1	79.69	79.70	67.42	67.39	78.35	78.28
2	89.41	89.45	67.18	67.09	82.65	82.59
3	91.34	91.28	75.21	75.63	89.95	89.87
4	81.27	81.09	67.78	67.81	85.26	85.28
5	70.98	71.13	66.85	66.79	74.73	74.80
6	99.89	98.04	77.93	77.96	90.65	90.81
7	76.02	76.19	67.95	67.85	75.35	75.41
8	78.77	78.85	68.22	68.16	74.53	74.68
9	87.60	87.59	66.96	66.83	89.78	89.64
10	79.83	79.93	76.85	77.06	83.59	83.62

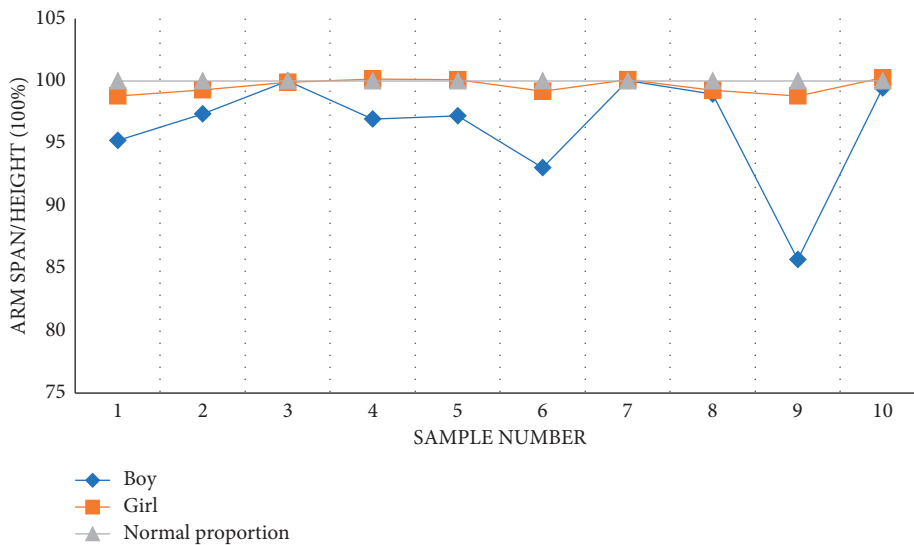


FIGURE 16: Basic index diagram of the body.

specific situation of the data more intuitively. As shown in Table 9, it is an analysis of the basic physical indicators of boys.

As shown in Figure 16, it is a broken line expression about the basic indicators of the body.

Analysis. By counting the basic physical indicators of all boys and girls (the data of 100 students are not shown here), we can find that the height range of boys is about (170, 180) and

that of girls is about (150, 165). On the whole, there is little difference in height, among which some are prominent in height and some are short in height. After calculation, we find that the height and arm length of the students are basically 1 : 1, which belongs to the normal range of ordinary people. In some cases, the arm span is greater than the height. According to the trend analysis of line chart, there is a positive correlation between students' arm span and height.

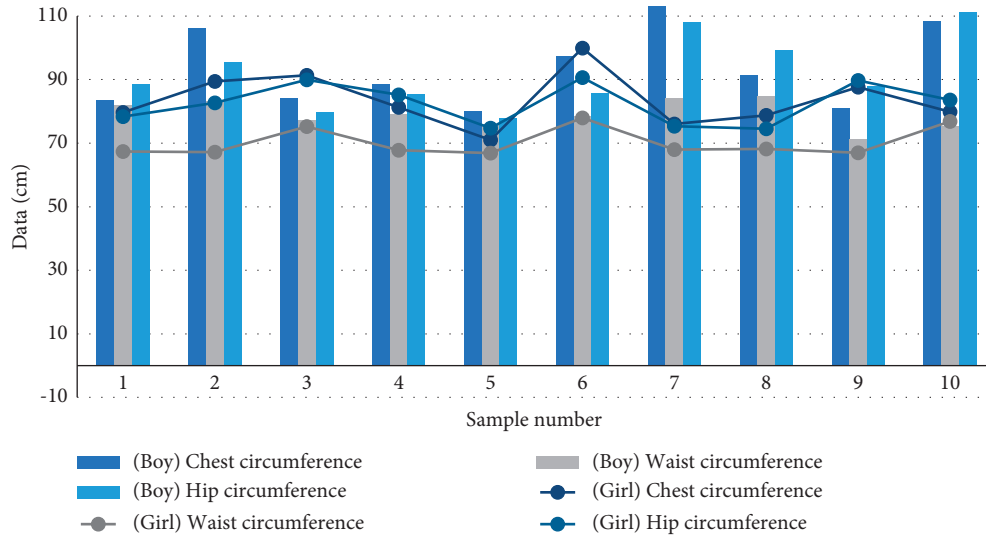


FIGURE 17: Body circumference index map.

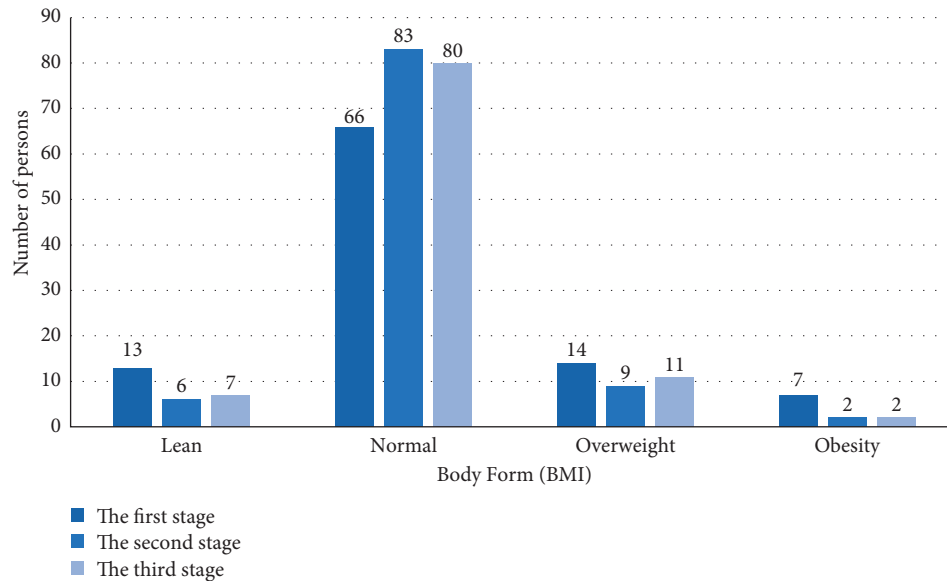


FIGURE 18: BMI exponential distribution in three stages.

(2) As shown in Figure 17, it is a broken line expression form about the body circumference index (part) of boys and girls.

Analysis. According to the above figure, we can find that no matter boys or girls, their individual circumference development is different, among which the chest circumference is the biggest difference, the hip circumference is the second, and the waist circumference curve is relatively gentle. On the whole, the fluctuation range of their respective circumference data curves is about 0 cm to 20 cm.

4.4.2. Comparative Analysis of Body Shape Indicators. In this part, we mainly compare three stages, BMI index, and vital capacity of male and female students.

(1) As shown in Figure 18, it is a comparative analysis chart of BMI index of all students in three stages.

As shown in Figure 19, it is a comparative analysis of the average BMI index of boys and girls.

Analysis. Before the first stage of training, students with abnormal body shape accounted for 34% of the total number, among which lean and overweight were the most. After two years of volleyball elective training, we can find that 83% of people have normal body shape, which is 25.76% higher than the previous number. The number of students who are thin, overweight, and obese has dropped significantly. After stopping training for one year, the number of students with normal shape decreased by about 3.61%, and the number of obese students did not increase again. We found that the weight of the subjects showed a downward trend as a whole.

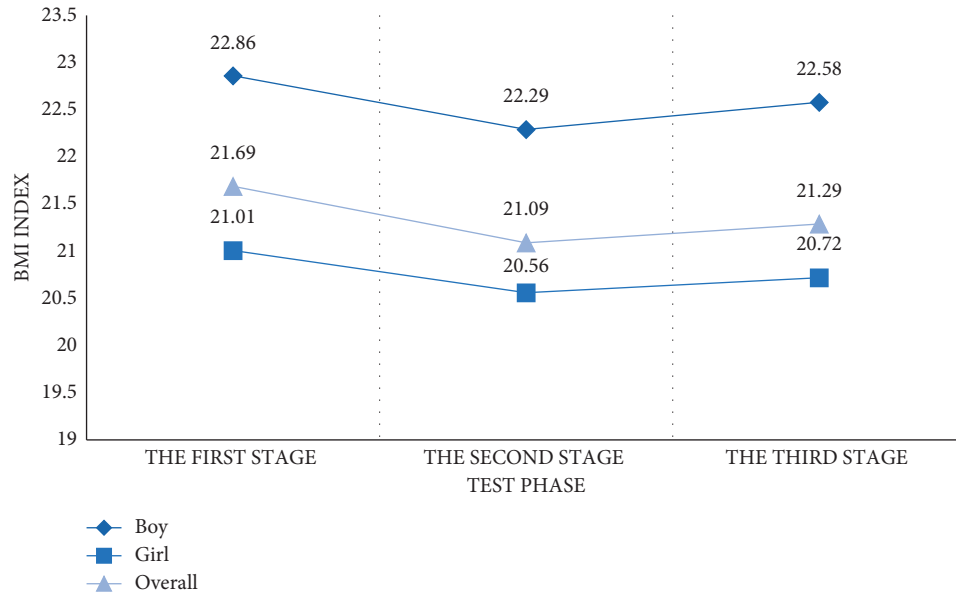


FIGURE 19: BMI mean distribution.

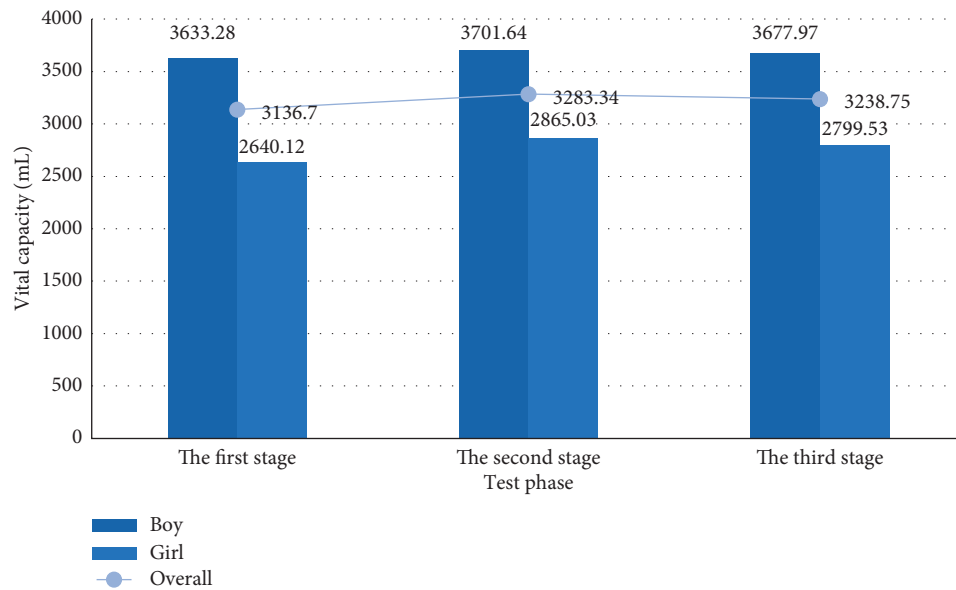


FIGURE 20: Distribution of mean vital capacity.

(2) As shown in Figure 20, it is about the average distribution of vital capacity of boys and girls.

Analysis. From the figure, we can find that the overall mean change of vital capacity of both boys and girls shows a small upward trend. In the second stage, the vital capacity of students increased the most. Although it decreased to a certain extent in the third stage, it can be seen that the cardiopulmonary respiratory system of students improved obviously.

4.4.3. Comparative Analysis of Physical Fitness Indicators. In this part, we divided the data into two tests, which are before and after the volleyball elective class. As shown in Tables 10 and 11, it is about the average analysis and comparison of physical fitness indicators of male and female students.

Analysis. According to Tables 10 and 11, we can find that no matter whether boys or girls, after training, the training effect of five items is obvious, and the physical fitness

TABLE 9: Basic body indicators (unit: cm).

Boy	Arm span/height (100%)	Girl	Arm span/height (100%)
1	95.25	1	98.80
2	97.36	2	99.30
3	100.01	3	99.90
4	96.95	4	100.16
5	97.22	5	100.11
6	93.05	6	99.20
7	100.04	7	100.12
8	98.95	8	99.27
9	85.71	9	98.80
10	99.44	10	100.26

TABLE 10: Comparison of physical fitness indexes of boys.

Project	Before training	After training
50 m	7.92 ± 6.77	7.31 ± 8.20
Standing long jump	216.30 ± 16.09	246.97 ± 13.12
Sitting body flexion	8.62 ± 3.3	8.35 ± 6.12
Pull-up	7.66 ± 4.03	10 ± 3.10
1000 m	238.10 ± 18.55	240.13 ± 19.01

TABLE 11: Comparison of physical fitness indexes of girls.

Project	Before training	After training
50 m	8.55 ± 10.22	8.10 ± 10.89
Standing long jump	165.05 ± 37.50	179.68 ± 41.07
Sitting body flexion	11.80 ± 6	10.2 ± 6.56
1 min sit-up	30.45 ± 7.55	37.52 ± 7.87
800 m	241.90 ± 13.30	241.41 ± 14.53

indicators of students have a certain increase. This shows that volleyball elective course can effectively help students improve their physical quality and sports ability.

5. Conclusion

Whether the change of human body constitution can grow healthily is a key issue for the state to pay attention to. The purpose of this study is to empirically analyze the impact of volleyball elective courses on students' physical health, to illustrate the role of volleyball elective courses by using theory and data, to prove that volleyball elective courses can effectively improve students' physical fitness, and to cultivate students' enthusiasm and interest in sports so as to guide students to love sports and match their physical fitness with their learning quality. This paper analyzes the influence of volleyball on students' physique by using digital image processing technology. Through the experiment, we can get the conclusion that after the system is adopted, the error of the measurement index is small. For boys and girls, the development of three circumferences is different, the difference of chest circumference is the largest, the difference of hip circumference is the second, and the curve of waist circumference is relatively gentle. After training, the body shape is larger than before, and the overall weight shows a downward trend.

Volleyball optional courses have different degrees of improvement on each student's physical fitness, and to achieve the best results, we need more diversified teaching experience. The focus of future work should be tailored according to their different physical qualities, and individualized teaching programs should be formulated and implemented in order to achieve the scientific goals proposed in the article.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this work.

References

- [1] S. Sunoj, S. N. Subhashree, S. Dharani et al., "Sunflower floral dimension measurements using digital image processing," *Computers and Electronics in Agriculture*, vol. 151, pp. 403–415, 2018.
- [2] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital image registration[J]," *IEEE Transactions on Computers*, vol. C-21, no. 2, pp. 179–186, 2009.
- [3] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: survey and analysis of current methods," *Signal Processing*, vol. 90, no. 3, pp. 727–752, 2010.
- [4] N. Nakano, A. Ohsumi, and K. Yoshida, "Identification of physical parameters of cantilevered beams from noisy vibration data," *Transactions of the Japan Society of Mechanical Engineers Series C*, vol. 67, no. 653, pp. 43–50, 2001.
- [5] I. Pitas, "Digital image processing algorithms and applications [J]," *IEEE Signal Processing Magazine*, vol. 18, no. 2, p. 58, 2000.
- [6] P. Subathra, M. Elango, and A. Subramani, "Combined effect of parcourse training and mental training on selected physical and psychological variables among volleyball players[J]," *Journal of Information and Computational Science*, vol. 11, no. 2, pp. 53–62, 2021.
- [7] P. Subathra, M. Elango, and A. Subramani, "Effect of par-course training on leg explosive power and shoulder strength among volleyball players[J]," *International Journal of Analytical and Experimental Modal Analysis*, vol. 13, no. 1, pp. 704–707, 2021.
- [8] S. Chen, Wanglili, and Z. Chen, "An empirical study on the application of stair -type teaching method in volleyball elective course in colleges and universities[J]," *Sports Science and Technology Literature Bulletin*, vol. 026, no. 006, pp. 91–93, 2018.
- [9] E. Kubota, T. Suzuki, M. Honda, and T. Ikenaga, "Action detection of volleyball using features based on clustering of body trajectories[J]," *The Journal of the Institute of Image Electronics Engineers of Japan*, vol. 45, no. 3, pp. 373–381, 2016.
- [10] G. Vasić, N. Trajković, D. Mačak et al., "Intensity-modified recreational volleyball training improves health markers and physical fitness in 25-55-year-old men," *BioMed Research International*, vol. 2021, no. 1, pp. 1–9, 2021.
- [11] J. Chen and S. Chen, "Study of the effect of four people's soft volleyball on the physical and mental health of college

- students[J],” *Journal of Jilin Institute of Physical Education*, vol. 022, no. 001, pp. 93-94, 2006.
- [12] D. K. Xiaofen, L. Harrison, L. Chen, P. Xiang, and J. Castro Piñero, “An analysis of research on student health-related fitness knowledge in K-16 physical education programs[J],” *Journal of Teaching in Physical Education*, vol. 28, no. 3, pp. 333-357, 2009.
- [13] S. Huddleston, J. Mertesdorf, and K. Araki, “Physical activity behavior and attitudes toward involvement among physical education, health, and leisure services pre-professionals[J],” *College Student Journal*, vol. 36, no. 1, pp. 111-117, 2002.
- [14] C. Konczos, J. Bognár, Z. Szakály, I. Barthalos, I. Simon, and Z. Oláh, “Health awareness, motor performance and physical activity of female university students,” *Biomedical Human Kinetics*, vol. 4, no. 2012, pp. 12-17, 2012.
- [15] T. Klein and S. Becker, “Age and exercise: a theoretical and empirical analysis of the effect of age and generation on physical activity,” *Journal of Public Health*, vol. 20, no. 1, pp. 11-21, 2012.
- [16] S. Dutta, S. K. Pal, S. Mukhopadhyay, and R. Sen, “Application of digital image processing in tool condition monitoring: a review,” *CIRP Journal of Manufacturing Science and Technology*, vol. 6, no. 3, pp. 212-232, 2013.
- [17] S. Yakin, T. Hasanuddin, and N. Kurniati, “Application of content based image retrieval in digital image search system,” *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 2, pp. 1122-1128, 2021.
- [18] G. B. Chen, Z. Sun, and L. Zhang, “Road identification algorithm for remote sensing images based on wavelet transform and recursive operator,” *IEEE Access*, vol. 8, pp. 141824-141837, 2020.
- [19] Y. Zhou and Y. Huang, “On the application of the practice method in the volleyball teaching in public sports course—taking the elective courses of public volleyball in a college in shaanxi province as an example.% the application of “replacing training with competition” in the teaching of public sports volleyball pad [J],” *Sports Boutique (Academic Edition)*, vol. 037, no. 007, pp. 18-20, 2018.
- [20] K. Y. Park, D. 김동원, Y. 어영숙, and B. 서보원, “Impact of sitting volleyball program on the isokinetic muscular strength and mental health of women with disabilities,” *Korean Journal of Physical, Multiple, & Health Disabilities*, vol. 53, no. 3, pp. 71-87, 2010.
- [21] S. Zhang and H. Mao, “Optimization analysis of tennis players’ physical fitness index based on data mining and mobile computing,” *Wireless Communications and Mobile Computing*, vol. 2021, no. 11, pp. 1-11, 2021.
- [22] X. Dai and S. Li, “Volleyball data analysis system and method based on machine learning,” *Wireless Communications and Mobile Computing*, vol. 2021, no. 20, pp. 1-11, 2021.
- [23] A. Watson, K. Biese, S. A. Kliethermes, E. Post, and T. McGuine, “Impact of in-season injury on quality of life and sleep duration in female youth volleyball athletes: a prospective study of 2073 players[J],” *British Journal of Sports Medicine*, vol. 55, no. 16, 2021.
- [24] H. Ji and C. Zheng, “The influence of physical exercise on college students’ mental health and social adaptability from the cognitive perspective[J],” *Work*, vol. 2021, no. 1, pp. 1-12, 2021.
- [25] M. J. Ormsbee, A. W. Kinsey, M. Chong, S. Heather, and Friedman, “The influence of high intensity interval training on the salivary cortisol response to a psychological stressor and mood state in non-sedentary college students[J],” *Journal of Exercise Physiology Online*, vol. 16, no. 1, 2013.

Research Article

A Data Augmentation Method for Prohibited Item X-Ray Pseudocolor Images in X-Ray Security Inspection Based on Wasserstein Generative Adversarial Network and Spatial-and-Channel Attention Block

Dongming Liu ^{1,2}, Jianchang Liu ^{1,2}, Peixin Yuan ³, and Feng Yu ^{1,2}

¹College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

²State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China

³School of Mechanical Engineering and Automation, Northeastern University, Shenyang 110819, China

Correspondence should be addressed to Jianchang Liu; liujianchang@mail.neu.edu.cn

Received 23 January 2022; Accepted 26 February 2022; Published 18 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Dongming 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.

For public security and crime prevention, the detection of prohibited items in X-ray security inspection based on deep learning has attracted widespread attention. However, the pseudocolor image dataset is scarce due to security, which brings an enormous challenge to the detection of prohibited items in X-ray security inspection. In this paper, a data augmentation method for prohibited item X-ray pseudocolor images in X-ray security inspection is proposed. Firstly, we design a framework of our method to achieve the dataset augmentation using the datasets with and without prohibited items. Secondly, in the framework, we design a spatial-and-channel attention block and a new base block to compose our X-ray Wasserstein generative adversarial network model with gradient penalty. The model directly generates high-quality dual-energy X-ray data instead of pseudocolor images. Thirdly, we design a composite strategy to composite the generated and real dual-energy X-ray data with background data into a new X-ray pseudocolor image, which can simulate the real overlapping relationship among items. Finally, two object detection models with and without our data augmentation method are applied to verify the effectiveness of our method. The experimental results demonstrate that our method can achieve the data augmentation for prohibited item X-ray pseudocolor images in X-ray security inspection effectively.

1. Introduction

With the frequent population flow, people carry more and more items in the baggage and the types of prohibited items have become increasingly abundant. In order to ensure public security, the research studies on the detection method of prohibited items are significant. Currently, X-ray inspection technology has been widely used in the security inspection of public places, reducing criminal behavior effectively. In the process of security inspection, security inspectors need to determine whether there are prohibited items in baggage through the X-ray images. In some special situations, such as the traffic rush hours, the frequency of baggage passing is greatly increased, which requires security inspectors to complete the inspection in a very short time. In

addition, there is still no unified standard for the training of security inspectors, and the accuracy of inspection depends on the experience and working status of security inspectors. For these reasons, the accuracy of manual detection methods cannot be guaranteed [1]. Accordingly, a fast and effective automatic detection method of prohibited items for X-ray security inspection is significant.

The detection method of prohibited items has been developed for many years. In the early stage, the automatic detection method of prohibited items usually used different feature extraction algorithms to extract the features of prohibited items and then classified the extracted features. These feature extraction methods included scale-invariant feature transform (SIFT), Haar-like features (Haar), bag of visual words (BOW), histogram of oriented gradients

(HOG), and so on [2]. Based on these feature extraction methods, researchers proposed some detection methods of prohibited items. In the case of the single-energy X-ray image, Turcsany et al. [3] proposed a method based on BOW for detecting firearms using a dual-view approach. In [4], a method using visual vocabulary and an occurrence structure was proposed to detect three different prohibited items. In addition, an approach called adaptive sparse representation (XASR+) [5] was proposed to recognize items automatically in cases with less constrained conditions including contrast variations, pose variations, image size variations, and focal distance variations. In the case of the dual-energy X-ray image, Riffó and Mery [6] proposed an active X-ray testing framework that is able to find an adequate view of the object item to detect razor blades in different cases. The applicability of BOW methods in X-ray image classification and retrieval was discussed in [7]. In practical applications, the performance of these automatic detection methods based on manual features cannot meet the requirements due to the wide variety of items, occlusion, noise, clutter, and other reasons.

In recent years, the convolutional neural networks (CNNs) have been widely used in image analysis and processing. Methods based on deep learning have achieved great success in many computer vision tasks [8–10]. A CNN-based classification algorithm was proposed in [9], after which CNN-based classifiers have received wide attention in the field of computer vision. These methods based on the CNN have achieved satisfactory results in image classification [11], object detection [12], target segmentation [13], etc. For the object detection methods based on the CNN, the methods are mainly divided into two classifications. One is the two-stage method, such as R-CNN [12], Fast R-CNN [14], Faster R-CNN [15], R-FCN [16], and FPN [17]. These methods first generate a set of candidate region suggestions and then classify, filter, and refine the candidate region suggestions to achieve object detection. The other is the one-stage method, such as YOLO [18–21] and SSD [22]. These methods predict the classification and bounding box directly from a single convolutional network. Such methods have a faster detection speed, but the accuracy is lower than that of the two-stage methods. Both of them have achieved success in the object detection of natural optical images and have been applied in various fields [23–25].

Compared with natural optical images, X-ray pseudocolor images are quite different in several aspects. Natural optical images are formed by the light reflection, while X-ray pseudocolor images are formed by irradiating the objects with X-ray, which loses a lot of information about the surface of objects [26]. In addition, X-ray pseudocolor images consist of shadows from overlapping transparent layers. The transparency of the pseudocolor image is determined by the material density along the X-ray path and different materials will appear in different colors. The overlap between objects also makes the same object appear in different colors, and high-density objects (e.g., thick metal) can obscure other overlapping objects [27]. These phenomena make the research on object detection of X-ray images difficult. In [28], Mery et al. used transfer learning to

classify three kinds of prohibited items based on the X-ray grayscale images, and the experimental results showed that the method is effective. In [29], Akcay et al. compared a BoVW approach with a CNN approach, and the experiments showed that the methods based on the CNN outperform the BoVW methods. In [30], the researchers proposed a method using deep convolutional neural networks to detect objects in X-ray security inspection. The method adopted a specific data enhancement technique, feature enhancement blocks, and multiscale fusion regions of interest. Most of the previous research studies for X-ray security inspection used transfer learning based on the ImageNet dataset. However, the direct adoption of the pretrained networks limits the adjustment of the network structure for X-ray pseudocolor images and may reduce the detection performance.

In addition, the method based on deep learning requires a large amount of data. Nevertheless, the X-ray pseudocolor image dataset is scarce and the color definition of X-ray pseudocolor images varies from company to company, which presents an enormous challenge to the object detection method of prohibited items for X-ray security inspection. There are two methods to solve this problem. One is to collect sufficient X-ray images containing the prohibited items with various poses and scales, which requires huge costs. The other is data augmentation. Using data augmentation methods can improve the generalization ability of the methods based on deep learning. Data augmentation methods include rotation, translation, scaling, etc. However, these methods have limited performance gains for the methods based on deep learning. In the last few years, the generative adversarial network (GAN) has achieved considerable success in image generation [31]. To improve the quality of the generated images, some models based on GAN have been proposed, such as the CGAN [32], the Cycle GAN [33], and the WGAN-GP [34]. Recently, some data augmentation methods based on the GAN have been used for X-ray image datasets. In [35], Yang et al. proposed a method of prohibited item X-ray pseudocolor image generation using the GAN. Their generated X-ray pseudocolor images only contain one single prohibited item and the image quality and diversity are not ideal. In [36], Zhu et al. proposed a method based on Cycle GAN to transform the item natural images into X-ray pseudocolor images. These methods used the methods based on GAN to directly generate prohibited item X-ray pseudocolor images and then replaced the target area of the real background X-ray pseudocolor image with the generated images. It is worth noting that these methods cannot show the real overlapping relationship among items. This also resulted in the fact that the final composite X-ray pseudocolor images are not ideal.

To generate more realistic and higher-quality X-ray pseudocolor images, we propose an effective data augmentation method for prohibited item X-ray pseudocolor images in X-ray security inspection (DA-PIX). The main contributions of the method are as follows. Firstly, we design the framework of the DA-PIX to achieve the dataset augmentation using the datasets with and without prohibited items. Secondly, based on the WGAN-GP, we design a

spatial-and-channel attention block (SCAB) and a new base block to compose our X-ray Wasserstein generative adversarial network model (SCAB-XWGAN-GP). The model directly generates high-quality dual-energy X-ray data instead of X-ray pseudocolor images. Thirdly, in order to generate more realistic X-ray pseudocolor images, we design a composite strategy based on the absorption law of X-ray to composite the generated and real dual-energy X-ray data with background data into a new more realistic X-ray pseudocolor image, which can simulate the real overlapping relationship among items.

The rest of the paper is organized as follows. Section 2 describes the dataset creation. In Section 3, the proposed method is described. Section 4 presents the experiments and results. Finally, Section 5 concludes the paper and discusses some directions for future work.

2. Dataset Creation

The dual-energy X-ray method has been widely used in X-ray security inspection systems. The method realizes the material classification of the detected objects by measuring the difference of attenuation coefficient of different materials under high-energy and low-energy X-ray. The low-energy and high-energy data are converted into an X-ray pseudocolor image by a lookup table to facilitate the interpretation of the detected objects, and the lookup table can be obtained through calibration [37]. In order to clearly distinguish the material of the detected object, orange represents the organic matter, green represents the mixture matter, and blue represents the inorganic matter in the X-ray pseudocolor image. Figure 1 shows the imaging process of the X-ray pseudocolor image.

The dual-energy X-ray security inspection equipment was provided by Shenyang DT Inspection Equipment Co., Ltd., China. The X-ray tube voltage is 140 kV. The X-ray tube current is 0.75 mA. The value range of the dual-energy X-ray data is normalized to 0–15200. The size of each data is $600 \times 600 \times 2$. Our X-ray dataset is divided into two parts. One is the X-ray prohibited item dataset (XD-P). The XD-P is collected in the simulated situation. To establish the XD-P, different baggage with the random prohibited items is packed and sent into the security inspection equipment. The XD-P consists of 8,000 dual-energy X-ray data and corresponding X-ray pseudocolor images. The prohibited items include brass knuckles, firecrackers, guns, hammers, knives, lighters, metal bottles, pliers, and scissors. The other is the X-ray security item dataset (XD-S). The data in the XD-S are the basic data for the data augmentation. The data in XD-S were collected from the real subway security inspection. The XD-S consists of 10,000 dual-energy X-ray data and corresponding X-ray pseudocolor images without prohibited items. Some X-ray pseudocolor images of the dual-energy X-ray data used in this paper are shown in Figure 2.

3. DA-PIX Method

The dataset plays an important role in the object detection methods based on deep learning. Due to the security, the

dataset is scarce. In order to augment the dataset and improve the robustness of the method, some methods based on the GAN are proposed. However, these methods are limited for our dataset because these methods cannot show the real overlapping relationship between the generated prohibited items and the background. This also results in the fact that the final composite X-ray pseudocolor images are not ideal. Accordingly, we propose an effective data augmentation method DA-PIX for prohibited item X-ray pseudocolor images in X-ray security inspection. Firstly, we design the framework of the DA-PIX. Secondly, we design the SCAB-XWGAN-GP model that directly generates high-quality dual-energy X-ray data instead of X-ray pseudocolor images. Thirdly, we design a composite strategy to composite the generated and real dual-energy X-ray data with background data into a new more realistic X-ray pseudocolor image.

3.1. The Framework of DA-PIX. To obtain more ideal composite X-ray pseudocolor images and improve the robustness of the detection method, we design the DA-PIX. The framework of the DA-PIX is shown in Figure 3. First, the SCAB-XWGAN-GP model is trained using a training dataset extracted manually from the XD-P. Second, the real dual-energy X-ray data of prohibited items and the generated dual-energy X-ray data of prohibited items using the SCAB-XWGAN-GP model are combined into a dataset for the composite strategy. Third, using the secure background dual-energy X-ray data from the XD-S and the foreground dual-energy X-ray data of prohibited items from the dataset for the composite dual-energy X-ray data, new composite dual-energy X-ray data can be obtained through the composite strategy. Then, the composite dual-energy X-ray data can be converted into an X-ray pseudocolor image by a lookup table. The composite images can more realistically simulate the actual situation. Finally, we can obtain the final augmented dataset.

3.2. SCAB-XWGAN-GP Model. Some methods based on GAN are applied to generate prohibited item X-ray pseudocolor images. A typical GAN model is composed of a generator and a discriminator. The architecture of the GAN model is shown in Figure 4. The generator is trained to generate the new data from random noise. The discriminator is trained to distinguish between the real data and the generated data.

However, these methods mainly focus on directly generating the X-ray pseudocolor images of prohibited items and cannot show the real overlapping relationship among items. Hence, we need to generate high-quality dual-energy X-ray data instead of X-ray pseudocolor images, which allows our next composite strategy to composite more realistic and higher-quality X-ray pseudocolor images. For the methods based on GAN, most models use upsampling + convolution or transposed convolution. Meanwhile, the X-ray prohibited items dataset is sparse and the training of the methods based on the GAN is challenging, which would result in the methods overfitting or collapse. In order to generate higher-quality prohibited item dual-energy

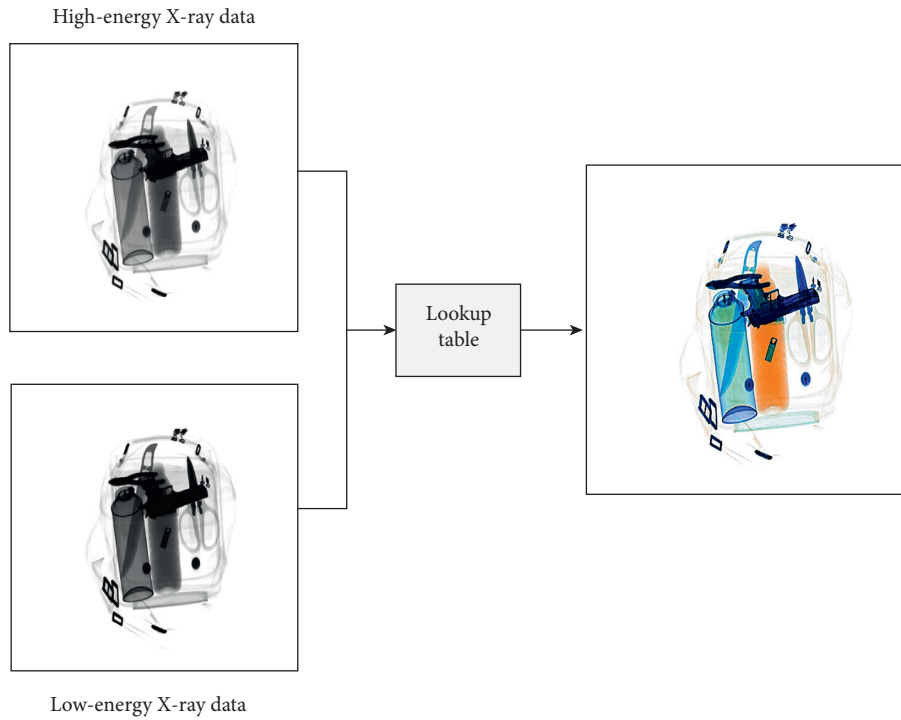


FIGURE 1: The imaging process of the X-ray pseudocolor image.

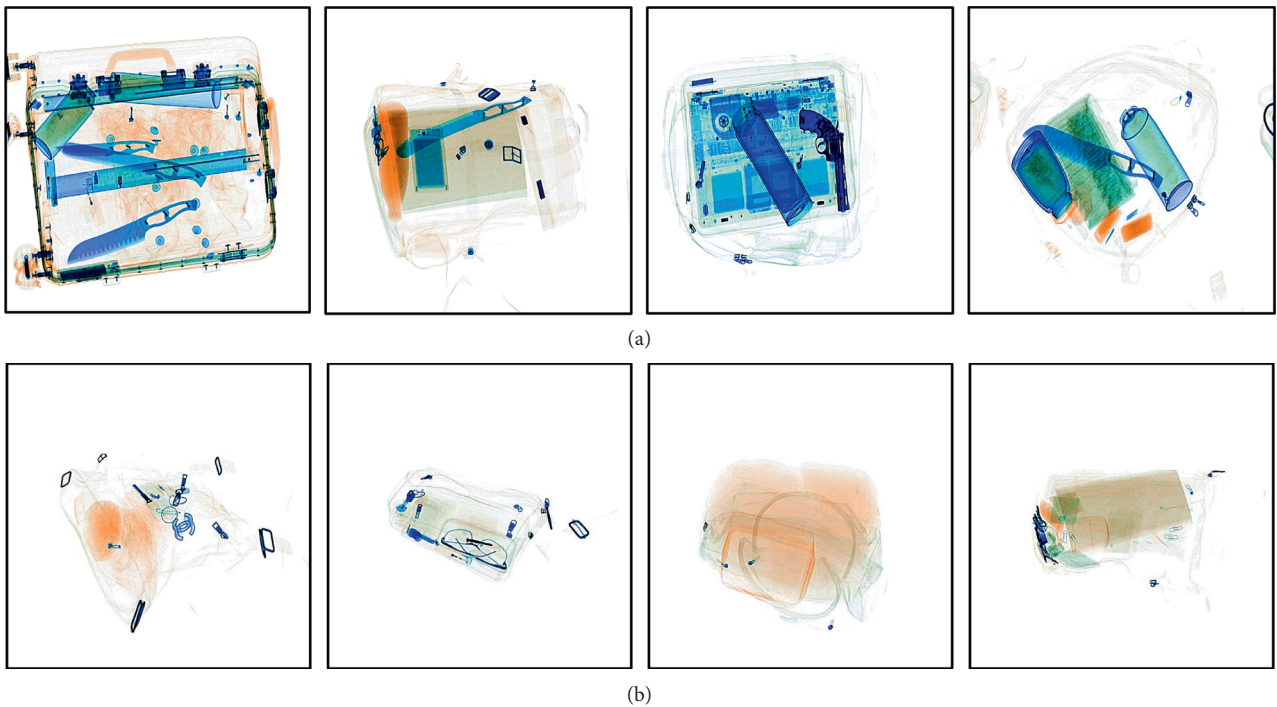


FIGURE 2: Some samples of our dataset. (a) Some samples of the XD-P. (b) Some samples of the XD-S.

X-ray data, we design the SCAB-XWGAN-GP model. In our model, we use the WGAN-GP as the basis and the dual-energy X-ray data as the generated object. Meanwhile, we design a new base block to compose the generator. The new base block can effectively improve the quality of the

generated images and reduce the training difficulty of the discriminator and the generator.

The structure of the base block is shown in Figure 5. In the base block, we employ the convolutional layer and the PixelShuffle in the subpixel convolution [38] to achieve

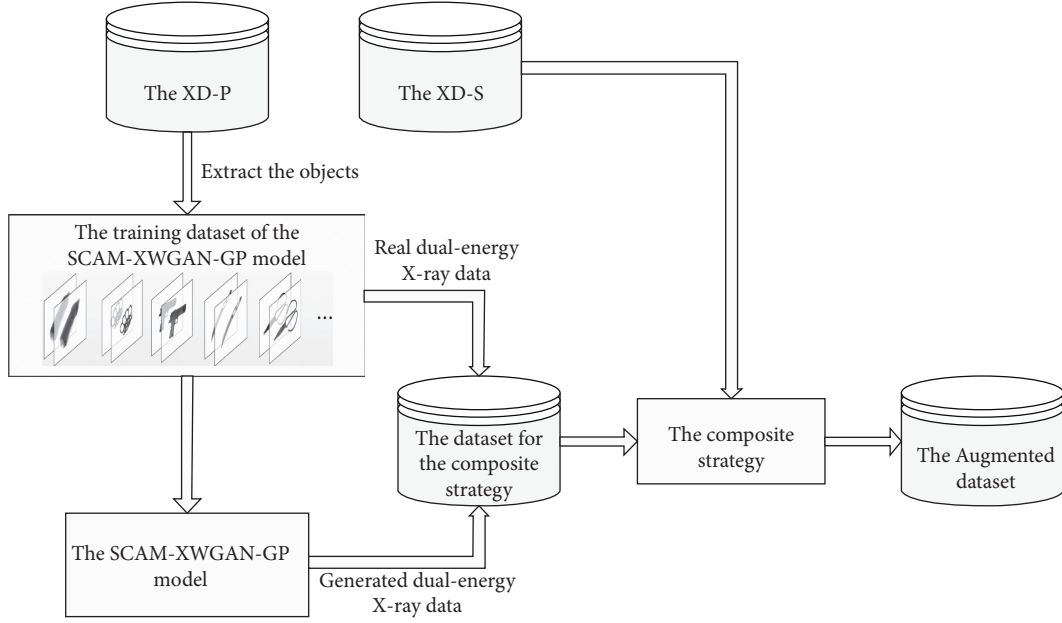


FIGURE 3: The framework of the DA-PIX.

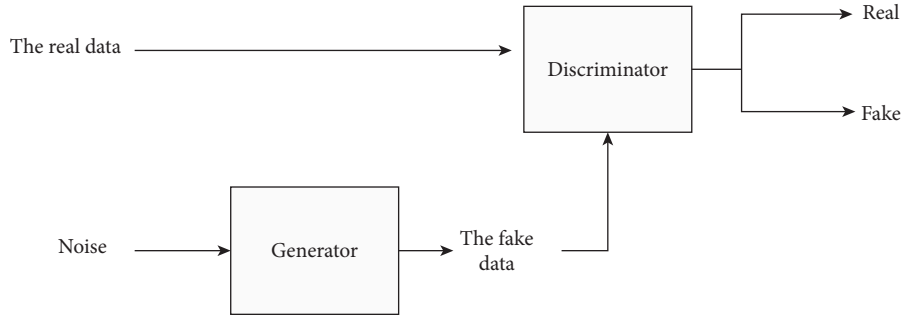


FIGURE 4: The architecture of the GAN model.

upsampling. The PixelShuffle can rearrange the elements of a $H \times W \times C \cdot r^2$ feature map to a feature map of shape $r \cdot H \times r \cdot W \times C$. After that, we use a residual block containing convolutional layers with 3×3 kernels to optimize the quality and detail of the generated data. In the process, in order to improve the fitting ability of the block and reduce the training difficulty, we use the ParametricRelu (Prelu) [39] as the activation function.

In addition, to generate more realistic dual-energy X-ray data, we design the SCAB based on the spatial-and-channel squeeze-and-excitation [40] block to fit our SCAB-XWGAN-GP model. The SCAB can make the details of data generated by the generator more realistic. Moreover, it can make the discriminator ignore less meaningful information and focus on more meaningful information. The architecture of the SCAB is shown in Figure 6. Given an input $X \in R^{H \times W \times C}$, we can get the calibrated output $\tilde{X} \in R^{H \times W \times C}$. The SCAB is divided into the spatial and channel branches.

For the channel branch, X can be expressed as $X = [x_1, x_2, \dots, x_p, \dots, x_C]$, where $x_i \in R^{H \times W}$. First, the

global average pooling and the global max pooling are used to generate two $1 \times 1 \times C$ feature maps $S_a = [S_{a1}, S_{a2}, \dots, S_{aC}]$ and $S_m = [S_{m1}, S_{m2}, \dots, S_{mC}]$ to express X in general. Second, the channel-wise dependencies \tilde{S} can be obtained using fully connected (FC) layers and nonlinearity layers. The process can be expressed as

$$\tilde{S} = \varphi(\mathbf{W}_{a2}\sigma(\mathbf{W}_{a1}S_a) + \mathbf{W}_{m2}\sigma(\mathbf{W}_{m1}S_m)), \quad (1)$$

where φ is the sigmoid function, σ represents the Relu activation function, $\mathbf{W}_{a1}, \mathbf{W}_{m1} \in R^{C/R \times C}$, $\mathbf{W}_{a2}, \mathbf{W}_{m2} \in R^{C \times C/r}$ are the weights of the fully connected layers, and r is a ratio parameter. Third, the recalibrated feature map of the channels \tilde{X}_c can be obtained by

$$\tilde{X}_c = \text{Scale}(\tilde{S}, X) = [\tilde{s}_1 x_1, \tilde{s}_2 x_2, \dots, \tilde{s}_C x_C], \quad (2)$$

where $\text{Scale}(\cdot)$ refers to channel-wise multiplication between the scalar \tilde{S} and the input X .

For the spatial branch, X can be expressed as $X = [x_{1,1}, x_{1,2}, \dots, x_{i,j}, \dots, x_{H,W}]$, where $x_{i,j} \in R^{1 \times 1 \times C}$. The spatial-wise dependencies O can be extracted using 1×1

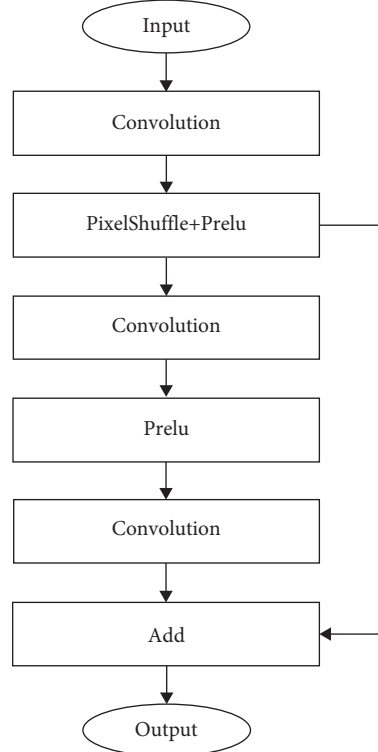


FIGURE 5: The structure of the base block.

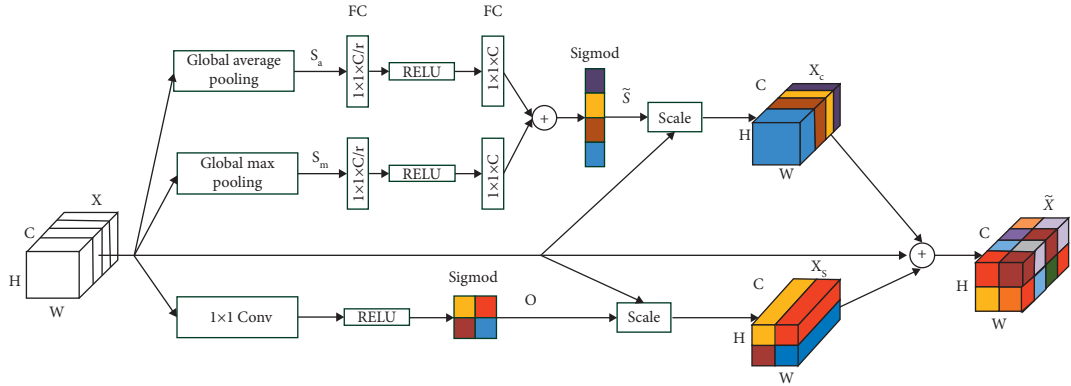


FIGURE 6: The architecture of the SCAB.

convolution layer and the nonlinearity layers. The process can be expressed as

$$O = \varphi(\sigma(\mathbf{W}_c X)), \quad (3)$$

where $\mathbf{W}_c \in R^{1 \times 1 \times C \times 1}$ is the weight of the 1×1 convolution layer. Then, the recalibrated feature map of the spatial can be obtained by

$$\tilde{X}_s = \text{Scale}(\tilde{S}, X) = [\tilde{s}_{1,1}x_{1,1}, \tilde{s}_{1,2}x_{1,2}, \dots, \tilde{s}_{H,W}x_{H,W}]. \quad (4)$$

After obtaining the recalibrated feature maps, the final calibrated output recalibrated feature map \tilde{X} can be obtained by

$$\tilde{X} = X + \tilde{X}_c + \tilde{X}_s. \quad (5)$$

Based on the base block and the SCAB, we design the SCAB-XWGAN-GP model. The architecture of the SCAB-XWGAN-GP model is shown in Figure 7. In the generator, six base blocks and one SCAB are concatenated to generate 256×256 data and the upsampling multiplier for each base block is 2. It is worth noting that the PixelShuffle requires four times the number of channels on its upper layer, which leads to a dramatic increase in computation and the number of parameters. Therefore, we limit the maximum number of channels of the base block to 256. It also means that the number of convolutional kernels per convolutional layer is at most 1024. Finally, we can obtain the generated dual-energy X-ray data through a convolution + tanh block. In the discriminator, the real and fake data are used as input data, and then the feature extraction is performed by six

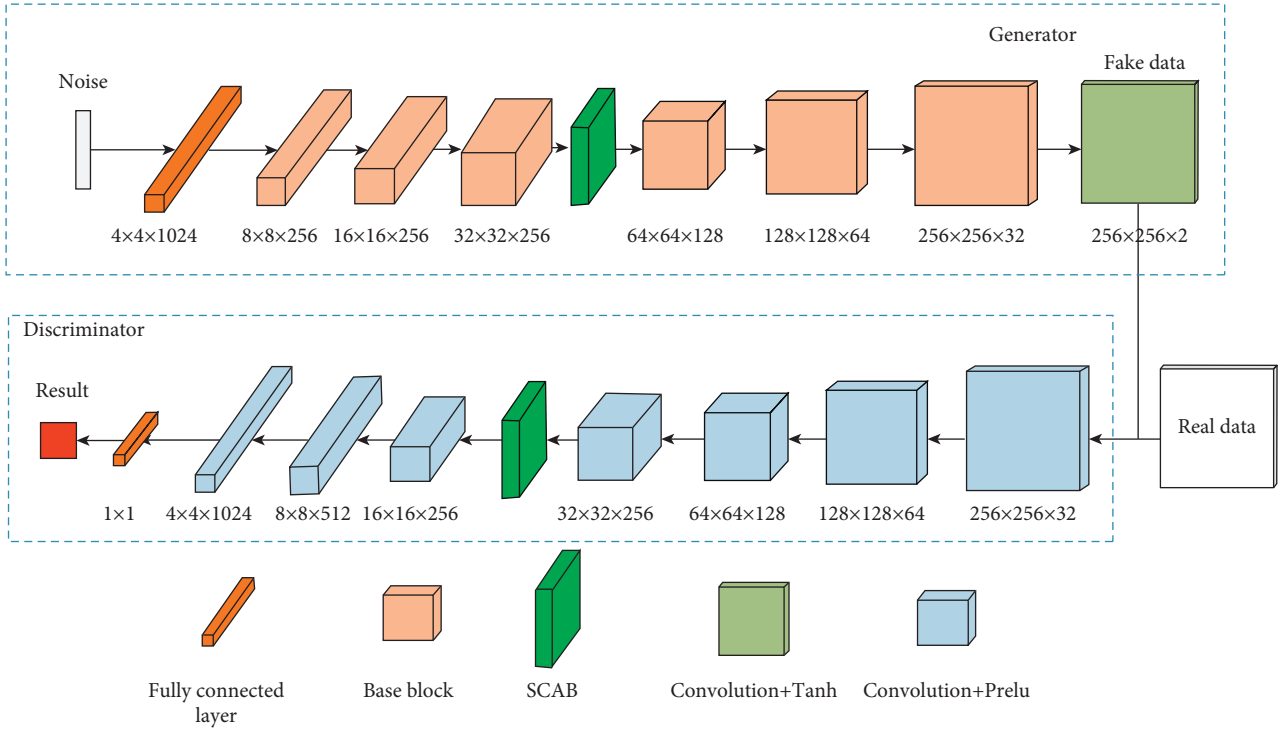


FIGURE 7: The architecture of the SCAB-XWGAN-GP (the number under each block represents the size of the data after passing through that block).

convolution + Prelu blocks and one SCAB. Finally, the predicted results are output through the fully connected layer.

For our SCAB-XWGAN-GP model, we use the same training method and the loss function as the WGAN-GP. The loss function L of the SCAB-XWGAN-GP model can be defined as

$$L = E[D(G(z))] - E[D(x)] + \lambda \cdot E\left[\left(\nabla_{\hat{x}} D(\hat{x})_2 - 1\right)^2\right], \quad (6)$$

where G is the generator, D is the discriminator, z is the random noise vector, λ is the penalty coefficient, and \hat{x} is the gradient penalty object sampled from the sample space of the generated data and the real data uniformly, $\hat{x} = \varepsilon z + (1 - \varepsilon)G(z)$, where $\varepsilon \in [0, 1]$.

3.3. The Composite Strategy. Through manual collection and the SCAB-XWGAN-GP model, we can obtain a large amount of dual-energy X-ray data for the prohibited items. For compositing the ideal X-ray pseudocolor image, we design the composite strategy based on the absorption law of X-ray. Based on the exponential law of photon radiation attenuation, the energy flux density after passage through the object with the thickness of d can be expressed as

$$\varphi = \varphi_0 e^{-\sigma d}, \quad (7)$$

where φ_0 is the incident energy flux density and σ is the attenuation coefficient. If X-ray radiation passes through n different objects, the energy flux density after passage through the objects can be expressed as

$$\varphi = \varphi_0 e^{\left(-\sum_{i=1}^n \sigma_i d_i\right)}, \quad (8)$$

where $i = 1, 2, \dots, n$. The gray value of a pixel can be linearly modeled as [37]

$$D = A \cdot \varphi + B, \quad (9)$$

where A and B are constant parameters of the model. Since the composite strategies for the high and low energies are same, we take single energy as an example. Define the gray value of the foreground object (D_{fore}) and the gray value of background object (D_{back}) as

$$D_{\text{fore}} = A \cdot \varphi_{\text{fore}} + B = A \cdot \varphi_0 e^{-\sigma_{\text{fore}} d_{\text{fore}}} + B, \quad (10)$$

$$D_{\text{back}} = A \cdot \varphi_{\text{back}} + B = A \cdot \varphi_0 e^{-\sigma_{\text{back}} d_{\text{back}}} + B, \quad (11)$$

where σ_{fore} and σ_{back} represent the attenuation coefficient of D_{fore} and D_{back} and d_{fore} and d_{back} represent the thickness of D_{fore} and D_{back} . From (8), the energy flux density of the composite data φ_c can be modeled as

$$\varphi_c = \varphi_0 e^{-\sigma_{\text{fore}} d_{\text{fore}} - \sigma_{\text{back}} d_{\text{back}}}. \quad (12)$$

The gray value of the composite data D_c can be expressed as

$$D_c = A \cdot \varphi_c + B = A \cdot \varphi_0 e^{-\sigma_{\text{fore}} d_{\text{fore}}} e^{-\sigma_{\text{back}} d_{\text{back}}} + B. \quad (13)$$

From (10), (11), and (13), we can obtain

$$\frac{D_c - B}{K} = \frac{D_{\text{fore}} - B}{K} \cdot \frac{D_{\text{back}} - B}{K}, \quad (14)$$

where K is the constant and it can be obtained using a calibration approach [37]. From (14), we can obtain the gray value of the composite data as follows:

$$D_c = \frac{(D_{\text{fore}} - B) \cdot (D_{\text{back}} - B)}{K} + B. \quad (15)$$

We can obtain a composite X-ray pseudocolor image using equation (15) and a lookup table. The details of the composite process are shown in Figure 8. Two random dual-energy X-ray data D_{fore} and D_{back} are selected (D_{back} belongs to the XD-S and D_{fore} belongs to the dataset for the composite images). Then, we randomly select the composite position P and generate the composite matrix $Mark$. $Mark$ is a two-valued matrix of zeros and ones, which is obtained from D_{fore} by binarization. Finally, the composite image can be obtained by

$$I_c = \text{Lookup}(D_{\text{back}} \cdot (1 - \text{Mark}) + C(D_{\text{back}}, D_{\text{fore}}, P, \text{Mark})), \quad (16)$$

where $\text{Lookup}(\cdot)$ represents the conversion of dual-energy X-ray data to X-ray pseudocolor images and $C(\cdot)$ indicates the composite operation of D_{back} and D_{fore} at position P using (15).

To demonstrate the superiority of our composite strategy, we show the original X-ray pseudocolor image and the composite X-ray pseudocolor images using different methods in Figure 9. The traditional method uses the foreground to replace the target area of the background. From Figure 9, we can find that the composite X-ray pseudocolor image using our composite strategy can simulate the real overlapping relationship between the prohibited item and the background compared to the traditional method, which makes the composite X-ray pseudocolor image more realistic.

In addition, our composite strategy can not only realize the composite of the single prohibited item and the safety images but also realize the composite of the multiple prohibited items and safety images or images containing prohibited items. Some composite X-ray pseudocolor images are shown in Figure 10. From Figure 10, we can find that the composite X-ray pseudocolor images using our method can simulate the real situation for the single prohibited item and multiple prohibited items. The composite X-ray pseudocolor images using our composite strategy can increase the diversity of our dataset.

Utilizing our composite strategy, we can randomly select the real and generated foreground objects to combine with the data in XD-S into an augmented dataset. By controlling the selection of the foreground objects and the composite position, we can easily control the position of objects in the augmented dataset and achieve a balance in the number of each object. Using the method, we can obtain a nearly infinite augmented dataset.

4. Experiments and Results

In this section, we first test our SCAB-XWGAN-GP model. Then, to verify the effectiveness of our DA-PIX, we design a

comparative experiment to evaluate the performance of our method. The experiments are run on a GPU system with the following specifications: Intel Core i9-10900k CPU, 64 GB RAM, and NVIDIA GeForce GTX 3090 GPU.

4.1. Evaluation Criteria. For evaluating the generated X-ray pseudocolor images, the Fréchet inception distance (FID) [41] is used. The FID is a comprehensive metric that has been shown to be more consistent with human assessments in assessing the realism and variability of the generated samples. We can obtain the FID by calculating the Wasserstein-2 distance between the generated X-ray pseudocolor images and the real X-ray pseudocolor images in the feature space of the Inception-v3 network. Lower FID value means that the distance between the generated data and the real data distribution is closer and the model works better.

For evaluating the performance of the object detection model with and without our DA-PIX, average precision (AP) and mean average precision (mAP) are used. AP can be calculated as

$$AP = \frac{1}{11} \sum_{Re \in \{0.0, 0.1, 0.2, \dots, 1\}} \max_{Re \geq Re} (Pr \tilde{Re}), \quad (17)$$

where Pr is the precision ($Pr = TP/TP + FP$), Re is the recall ($Re = TP/TP + FN$), TP is the number of true-positive samples, FP is the number of false-positive samples, FN is the number of false-negative samples, and $Pr(Re)$ is the measured precision at recall Re . Subsequently, mAP can be defined as

$$mAP = \frac{1}{K} \sum_{k=1}^K AP_k. \quad (18)$$

4.2. Results and Discussion. In order to better verify the effectiveness of our SCAB-XWGAN-GP model, we trained the DCGAN model and the WGAN-GP model to compare them with our model. For the training dataset used by the generative models, first we extracted the object data from the labeled XD-P by labeling the extracted data manually. Second, we kept the target data unchanged and replaced the background with 15200 by manual labeling. Third, we filled these data into a square with the maximum value of the width and height of the data and resized them to $256 \times 256 \times 2$ using the bilinear interpolation. Finally, these dual-energy X-ray data formed the training dataset used by the generative models. The adaptive moment estimation (Adam) optimization algorithm was used. The batch size was 32, the penalty coefficient λ was 10, and the learning rate of the generator and discriminator was 0.0002 and 0.0001, respectively. The update ratio of generator and discriminator was 5 for our model and the WGAN-GP model. For the DCGAN model, the update ratio of generator and discriminator was 1. It is worth noting that the generated data by these models are dual-energy X-ray data, and these data are converted into the X-ray pseudocolor images by a lookup

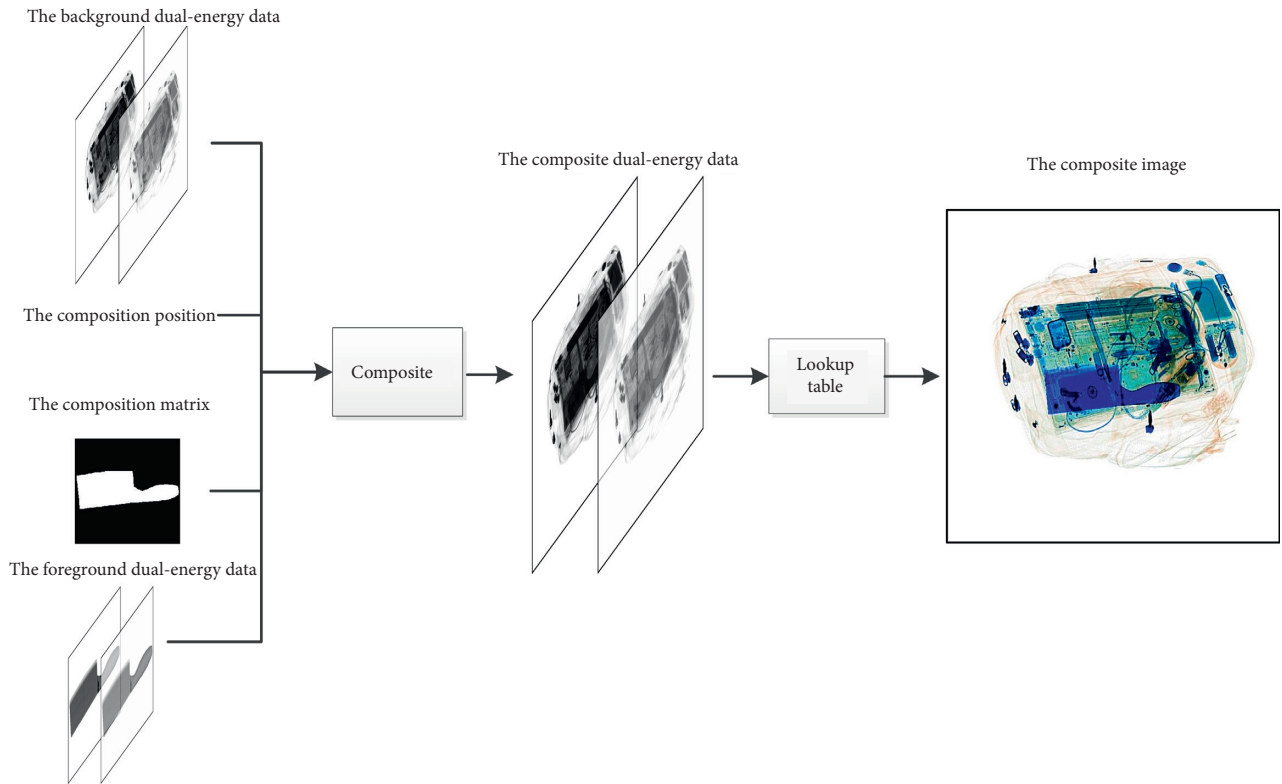


FIGURE 8: The composite process.

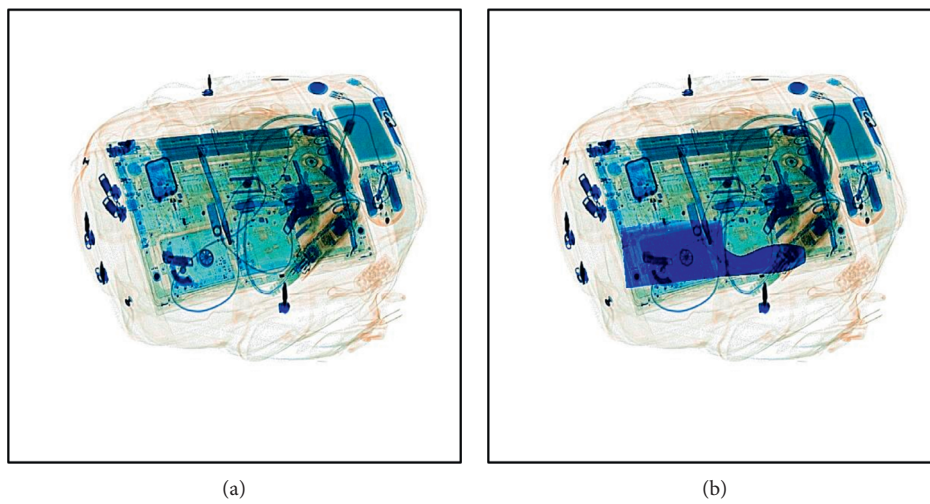
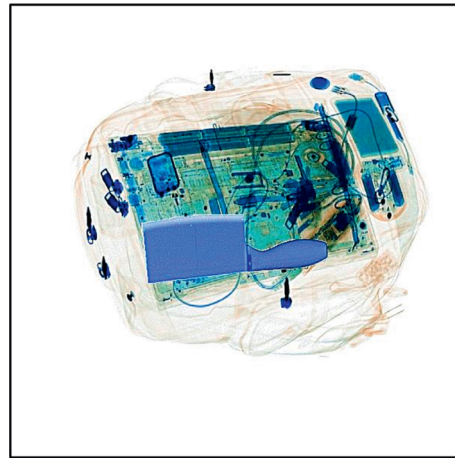
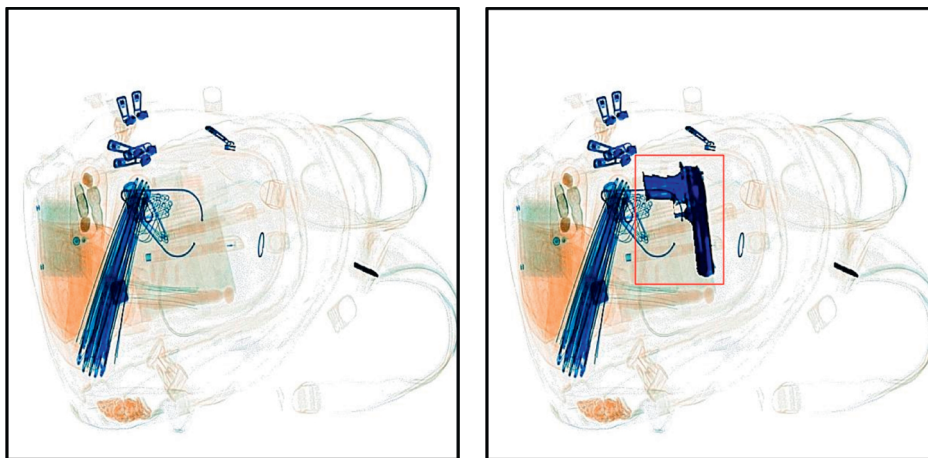


FIGURE 9: Continued.



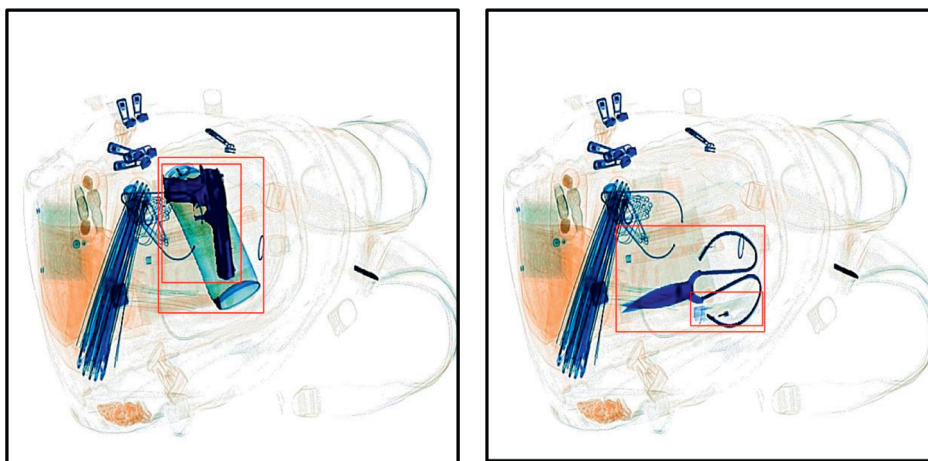
(c)

FIGURE 9: The original X-ray pseudocolor image and the composite X-ray pseudocolor images using different methods. (a) The original X-ray pseudocolor image. (b) The composite X-ray pseudocolor image using our composite strategy. (c) The composite X-ray pseudocolor image using the traditional method.



(a)

(b)



(c)

(d)

FIGURE 10: Continued.

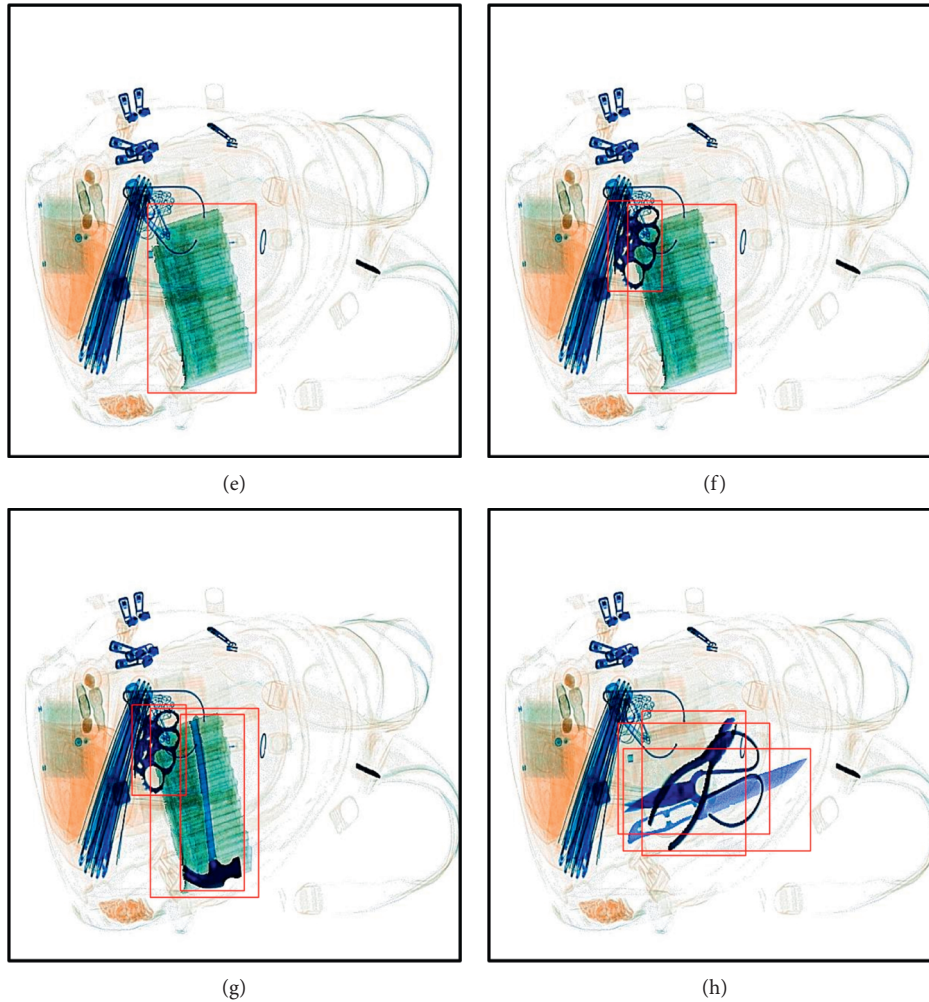


FIGURE 10: Some composite X-ray pseudocolor images. (a) The original X-ray pseudocolor image. (b–h) The composite X-ray pseudocolor images.

table. Some X-ray pseudocolor images generated by these models are shown in Figure 11.

From Figure 11, it can be found that the generated X-ray pseudocolor images by the DCGAN have a lot of noise, the images appear to have severe distortion, and the detailed information is blurred. Meanwhile, for the brass knuckle and plier, the DCGAN showed pattern collapse. For the WGAN-GP, the quality of the generated X-ray pseudocolor images has improved, but there is still noise and distortion in the images, the detailed information is still blurred, and the visual quality is poor. Compared with the methods, although there is still a little noise and minor deformation in the generated X-ray pseudocolor images, the quality of the X-ray pseudocolor images generated using our SCAB-XWGAN-GP is significantly improved.

In addition, to objectively evaluate our method, we compared the methods quantitatively using the FID score. The FID scores of these models are shown in Table 1. From Table 1, we can find that our method has the lowest FID score. This means that our model can better approximate the distribution of the real dual-energy X-ray data of the prohibited items.

Then, to verify the effectiveness of our DA-PIX, we trained the object detection models with and without our DA-PIX. We employed the YOLOV4-tiny model [21] as the object detection model. We randomly divided the XD-P into the training dataset and the testing dataset. The training dataset contained 7200 X-ray pseudocolor images and the test dataset contained 800 X-ray pseudocolor images. Our DA-PIX was used to augment the training dataset. In this process, first, we manually cropped the generated dual-energy X-ray data and resized them to their real size. Second, we combined these dual-energy X-ray data with the real dual-energy X-ray data extracted from the XD-P to form the dataset for the composite strategy. Third, we used the composite strategy to composite the new X-ray pseudocolor images with the data in the XD-S as the backgrounds. These composite X-ray pseudocolor images contained 2–7 prohibited items in each image. The augmented training dataset contained 17,200 X-ray pseudocolor images. After augmenting the dataset, we trained the YOLOV4-tiny model on the training dataset and the augmented training dataset, respectively. The stochastic gradient descent with momentum (SGDM) was used. The batch size was 16, the initial

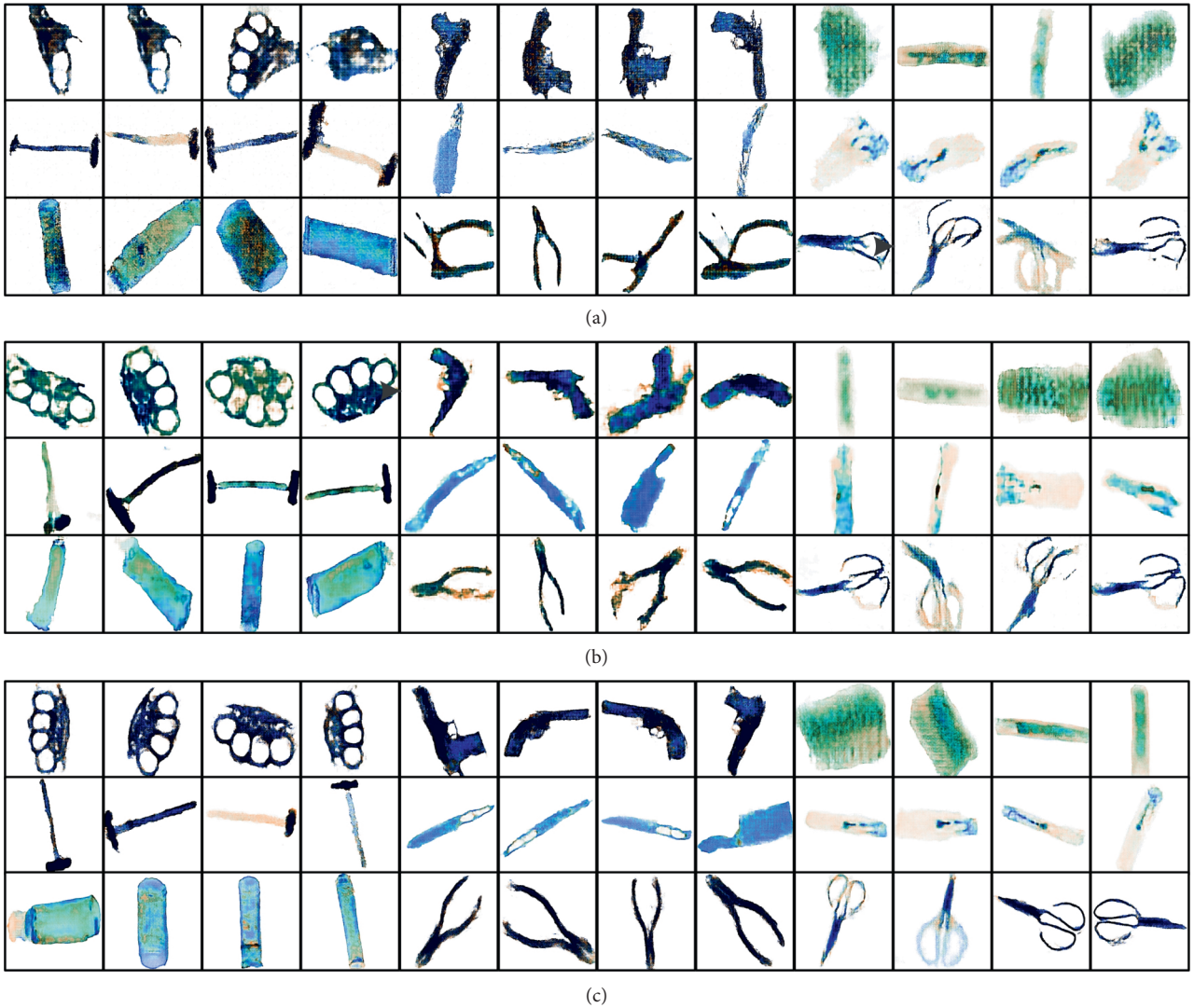


FIGURE 11: Some X-ray pseudocolor images generated by different models. (a) The X-ray pseudocolor images generated by the DCGAN. (b) The X-ray pseudocolor images generated by the WGAN-GP. (c) The X-ray pseudocolor images generated by the SCAB-XWGAN-GP.

TABLE 1: The FID scores of different models.

Items	WGAN-GP	DCGAN	SCAB-XWGAN-GP
Brass knuckle	185.3	143.7	103.2.
Firecracker	16.3	18.7	13.8
Gun	73.1	121.7	18.47
Hammer	32.6	41.8	19.2
Knife	25.8	32.3	9.4
Lighter	106.9	108.6	78.3
Metal bottle	45.2	48.6	41.4
Plier	43.8	45.6	27.2
Scissors	44.5	48.4	22.4

learning rate with warm-up was 0.001, and the momentum was 0.9. The detection results are shown in Table 2.

From Table 2, we can find that the mAP of the YOLOV4-tiny model trained by the augmented training dataset is higher than that of the YOLOV4-tiny model trained by the original training dataset, with an increase of 9.25%.

Moreover, the AP of each prohibited item has been enhanced using our method. It means that our DA-PIX method can effectively improve the performance of the detection model for prohibited item X-ray pseudocolor images in X-ray security inspection. In addition, the YOLOV4-tiny model is a lightweight detection model, and

TABLE 2: The detection results.

Items	YOLOV4-tiny + the original training dataset	YOLOV4-tiny + the augmented training dataset
Brass knuckle	88.75%	95.76%
Firecracker	86.63%	92.88%
Gun	86.80%	93.84%
Hammer	67.81%	80.59%
Knife	74.20%	80.58%
Lighter	86.69%	94.81%
Mental bottle	93.17%	96.23%
Plier	89.77%	90.11%
Scissors	72.79%	87.06%
mAP	82.96%	90.21%

we believe that the mAP will be further improved under more complex detection models.

5. Conclusion

In this paper, an effective data augmentation method for prohibited item X-ray pseudocolor images in X-ray security inspection is proposed. The innovation is mainly reflected in three major aspects. First, we design the framework of our DA-PIX to achieve dataset augmentation using the datasets with and without prohibited items. Second, in order to generate high-quality dual-energy X-ray data, a SCAB-XWGAN-GP model is designed based on the SCAB and a new base block. Third, a composite strategy based on the absorption law of X-ray is proposed to composite a new X-ray pseudocolor image, which can simulate the real overlapping relationship among items. Two YOLOV4-tiny models with and without our DA-PIX are trained to verify the effectiveness of our method. The experimental results demonstrate that our DA-PIX can effectively improve the mAP of the YOLOV4-tiny model. Therefore, our DA-PIX is effective for the data augmentation of prohibited item X-ray pseudocolor images in X-ray security inspection.

In future work, we will focus on designing effective detection models for the detection of prohibited item in X-ray security inspection.

Data Availability

The dataset used to support the findings of this study was supplied by Shenyang DT Inspection Equipment Co., Ltd., China, under license and the dataset involving security cannot be shared.

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 National Natural Science Foundation of China (no. 61773106) and Major Scientific and Technological Projects of the Ministry (no. JB2016GD034).

References

- [1] S. Michel, J. Ruiter, M. Hogervorst, S. Koller, and A. Schwaninger, "Computer-based training increases efficiency in x-ray image interpretation by aviation security screeners," in *Proceedings of the 41st Annual IEEE International Carnahan Conference on Security Technology*, pp. 201–206, Ottawa, ON, Canada, 8–11 Oct. 2007.
- [2] V. Rizzo, S. Flores, and D. Mery, "Threat objects detection in x-ray images using an active vision approach," *Journal of Nondestructive Evaluation*, vol. 36, no. 44, 2017.
- [3] D. Turcsany, A. Mouton, and T. P. Breckon, "Improving feature-based object recognition for x-ray baggage security screening using primed visual words," in *Proceedings of the 2013 IEEE International Conference on Industrial Technology (ICIT)*, pp. 1140–1145, Cape Town, South Africa, 25–28 Feb. 2013.
- [4] V. Rizzo and D. Mery, "Automated detection of threat objects using adapted implicit shape model," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 4, pp. 472–482, 2016.
- [5] D. Mery, E. Svec, and M. Arias, "Object recognition in x-ray testing using adaptive sparse representations," *Journal of Nondestructive Evaluation*, vol. 35, no. 3, pp. 1–9, 2016.
- [6] V. Rizzo and D. Mery, "Active x-ray testing of complex objects," *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 54, no. 1, pp. 28–35, 2012.
- [7] M. Bastan, M. R. Yousefi, and T. M. Breuel, "Visual words on baggage x-ray images," in *Proceedings of the Computer Analysis of Images and Patterns*, pp. 360–368, Seville, Spain, August 2011.
- [8] A. S. Elkorany and Z. F. Elsharkawy, "Covidetection-net: A tailored covid-19 detection from chest radiography images using deep learning," *Optik*, vol. 231, 2021.
- [9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [10] O. Russakovsky, J. Deng, H. Su et al., "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, 27–30 June 2016.
- [12] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 142–158, 2016.
- [13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic

- segmentation,” in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, 23–28 June 2014.
- [14] R. Girshick, “Fast r-cnn,” in *Proceedings of the 2015 IEEE International Conference on Computer Vision*, pp. 1440–1448, Santiago, Chile, 7–13 Dec. 2015.
- [15] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [16] J. Dai, Y. Li, K. He, and J. Sun, “R-fcn: object detection via regionbased fully convolutional networks,” in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 379–387, Barcelona Spain, December 2016.
- [17] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature pyramid networks for object detection,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 936–944, Honolulu, HI, USA, 21–26 July 2017.
- [18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: unified, real-time object detection,” in *Proceedings of the 2016 IEEE conference on computer vision and pattern recognition*, pp. 779–788, Las Vegas, NV, USA, 27–30 June 2016.
- [19] J. Redmon and A. Farhadi, “Yolo9000: better, faster, stronger,” in *Proceedings of the 2017 IEEE conference on computer vision and pattern recognition*, pp. 6517–6525, Honolulu, HI, USA, 21–26 July 2017.
- [20] J. Redmon and A. Farhadi, “Yolov3: an incremental improvement,” 2018, <https://arxiv.org/abs/1804.02767>.
- [21] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, “Yolov4: optimal speed and accuracy of object detection,” 2020, <https://arxiv.org/abs/2004.10934>.
- [22] W. Liu, D. Anguelov, D. Erhan et al., “Ssd: single shot multibox detector,” in *Proceedings of the European Conference on Computer Vision*, Amsterdam, The Netherlands, October 2016.
- [23] Z. Zhao, K. Hao, X. Ma et al., “SAI-YOLO: a lightweight network for real-time detection of driver mask-wearing specification on resource-constrained devices,” *Computational Intelligence and Neuroscience*, vol. 2021, p. 15, Article ID 4529107, 2021.
- [24] Z. Wu, X. Wang, and C. Chen, “Research on lightweight infrared pedestrian detection model algorithm for embedded platform,” *Security and Communication Networks*, vol. 2021, p. 7, Article ID 1549772, 2021.
- [25] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, “Fighting against covid-19: a novel deep learning model based on yolo-v2 with resnet-50 for medical face mask detection,” *Sustainable Cities and Society*, vol. 65, 2021.
- [26] Z. Chen, Y. Zheng, B. Abidi, D. L. Page, and M. Abidi, “A combinational approach to the fusion, de-noising and enhancement of dual-energy x-ray luggage images,” in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05) - Workshops*, San Diego, CA, USA, 21–23 Sept. 2005.
- [27] V. Rebuffel and J.-M. Dinten, “Dual-energy x-ray imaging: benefits and limits,” *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 49, no. 10, pp. 589–594, 2007.
- [28] D. Mery, E. Svec, M. Arias, V. Rizzo, J. M. Saavedra, and S. Banerjee, “Modern computer vision techniques for x-ray testing in baggage inspection,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 4, pp. 682–692, 2017.
- [29] S. Akçay, M. E. Kundegorski, M. Devereux, and T. P. Breckon, “Transfer learning using convolutional neural networks for object classification within x-ray baggage security imagery,” in *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*, pp. 1057–1061, Phoenix, AZ, USA, 25–28 Sept. 2016.
- [30] B. Gu, R. Ge, Y. Chen, L. Luo, and G. Coatrieux, “Automatic and robust object detection in x-ray baggage inspection using deep convolutional neural networks,” *IEEE Transactions on Industrial Electronics*, vol. 68, no. 10, pp. 10248–10257, 2021.
- [31] C. Ledig, L. Theis, F. Huzár et al., “Photo-realistic single image super-resolution using a generative adversarial network,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 105–114, Honolulu, HI, USA, 21–26 July 2017.
- [32] M. Mirza and S. Osindero, “Conditional generative adversarial nets,” 2014, <https://arxiv.org/abs/1411.1784>.
- [33] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2242–2251, Venice, Italy, 22–29 Oct. 2017.
- [34] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, “Improved training of wasserstein gans,” 2017, <https://arxiv.org/abs/1704.00028>.
- [35] J. Yang, Z. Zhao, H. Zhang, and Y. Shi, “Data augmentation for x-ray prohibited item images using generative adversarial networks,” *IEEE Access*, vol. 7, pp. 28894–28902, 2019.
- [36] Y. Zhu, Y. Zhang, H. Zhang, J. Yang, and Z. Zhao, “Data augmentation of x-ray images in baggage inspection based on generative adversarial networks,” *IEEE Access*, vol. 8, pp. 86536–86544, 2020.
- [37] D. Mery, *Computer Vision for X-ray Testing*, Springer International Publishing, Heidelberg, Germany, 2015.
- [38] W. Shi, J. Caballero, F. Huzár et al., “Real-Time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1874–1883, Las Vegas, NV, USA, 27–30 June 2016.
- [39] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: surpassing human-level performance on ImageNet classification,” in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1026–1034, Santiago, Chile, 7–13 Dec. 2015.
- [40] A. G. Roy, N. Navab, and C. Wachinger, “Concurrent spatial and channel ‘Squeeze & excitation’ in fully convolutional networks,” *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018*, vol. 11070, pp. 421–429, 2018.
- [41] H. Heuse, H. Ramsauer, B. Unterthiner, B. Nessler, and S. Hochreiter, “GANs trained by a two time-scale update rule converge to a local Nash equilibrium,” in *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 6629–6640, Montreal Canada, December 2017.

Research Article

Prediction Model of Stress Intensity Factor of Circumferential Through Crack in Elbow Based on Neural Network

Xiaohong Li ^{1,2}, Xianghui Li,¹ and Bin Chen¹

¹Liaoning Petrochemical University, Fushun, Liaoning 113001, China

²China University of Petroleum (East China), Qingdao, Shandong 266000, China

Correspondence should be addressed to Xiaohong Li; lixiaohong@lnpu.edu.cn

Received 11 January 2022; Revised 6 February 2022; Accepted 19 February 2022; Published 17 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiaohong 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.

Using ANSYS software to establish the finite element model of crack bending tube, the SIF at the tip of the crack is calculated for the difference in the diameter of the pipe, the outer diameter of the elbow, and the bending angle of the bend pipe, and it is used as a neural network to calculate the sample. By using three layers of BP network to establish the prediction model of the SIF of cracked pipe, the simulation of 39 sets of samples proves that the relative error of the BP network model is 0.19% and the mean square error of the network output is 0.0102. The prediction model has high prediction precision and generalization ability and can be used in engineering design and calculation.

1. Introduction

Elbows are widely used in engineering structural parts such as building structures, mechanical equipment, and vehicle manufacturing. Defects and cracks are usually inevitable. Tension and bending moment are the load types that pipes bear more. Therefore, it is of great significance to study the fracture parameters of circumferential cracks on the inner wall of elbows under tension and bending moment for the safety assessment of defective structures [1].

Stress intensity factor is an important parameter to control the fracture structure. It is a necessary theoretical basis to determine the service life of crack structure and design crack prevention measures. Fett and Noda [2, 3] used the analytical method to solve the strength factor, but it is only suitable for the relatively simple crack model. Rong et al. [4] recorded a variety of typical crack strength factor calculation formulas, but their mechanical models are infinite and semi-infinite plate structures and bear a single load. For more complex cracks, the classical theory is difficult to give the calculation method of

stress intensity factor, which is often solved by numerical methods such as the finite element method, finite difference method, and boundary element method [5]. The finite element method [6] is simple and accurate and has been proved to be an effective method for calculating the stress intensity factor of structures with cracks [7, 8]. Many parameters need to be considered in the strength factor of cracked elbow.

Due to the high complexity and professionalism of the operation of the finite element method, the calculator cannot complete all the calculations of a large number of variables. In this paper, the strength factor training samples are collected by the finite element method, and the artificial neural network method is used. The artificial neural network has strong memory and prediction ability to establish the calculation model of the stress intensity factor of the circumferential through crack of the elbow, so as to realize the calculation of the stress intensity factor of the circumferential through crack of the elbow in a certain range. The results show that the accuracy of the calculation network meets the engineering requirements.

2. Calculation of Stress Intensity Factor

The relationship between the stress intensity factor near the crack tip and stress, displacement, and strain is shown in formulas (1)–(3):

$$\sigma_{ij}^{(N)} = \frac{K_N}{\sqrt{2\pi r}} f_{ij}^{(N)}(\theta), \quad (1)$$

$$U_i^N = K_N \sqrt{\frac{r}{\pi}} g_i^{(N)}(\theta), \quad (2)$$

$$\varepsilon_{ij}^{(N)} = \frac{K_N}{\sqrt{2\pi r}} h_{ij}^{(N)}(\theta), \quad (3)$$

where $\sigma_{ij}(i, j=1, 2, 3)$ represents the stress component, $u_i(i=1, 2, 3)$ represents the displacement component, $N=I, II, III$ represents the type of fracture crack, and r and θ are distributed as the polar radius and rotation angle of the plastic zone at the crack tip.

K_I is used to represent the intensity factor of the stress field at the tip of I crack. Generally, it is considered that the calculation formula of I crack intensity factor is

$$K_I = \lim_{r \rightarrow 0} \sqrt{2\pi r} \sigma_y |_{\theta=0}. \quad (4)$$

It can be seen that the stress distribution near the crack tip is a function of r and θ , which is independent of the load borne by the material or structure and the crack length. The general expression of stress intensity factor is

$$K_I = Y\sigma\sqrt{\pi a}, \quad (5)$$

where σ is the nominal stress; a is the crack size; and Y is the shape factor. Literature review shows that the influencing factors of crack bend shape coefficient Y mainly include bend inner diameter, bend outer diameter, bend angle, and so on [9, 10].

3. Finite Element Modeling of Circular Through Crack Elbow

3.1. Structure and Geometric Parameters of Circumferential Through Crack Elbow. The common through elliptical crack elbow is selected as the analysis object. The geometric configuration and load of the crack in the elbow area are shown in Figure 1. Figure 2 shows the structural dimension diagram of the elbow, in which a is the inner diameter of the elbow, b is the outer diameter of the elbow, x is the center angle of the crack, α is the bending angle of the elbow, t is the wall thickness of the elbow, and D is the diameter of the elbow orifice.

3.2. Finite Element Modeling. To solve the crack stress intensity factor K by the finite element method, the three-dimensional model with crack must be established first. According to the characteristics of ANSYS finite element analysis [11], this paper establishes a symmetrical half bend model and completes the finite element meshing, as shown in Figure 3. The selected element is the shell element shell36

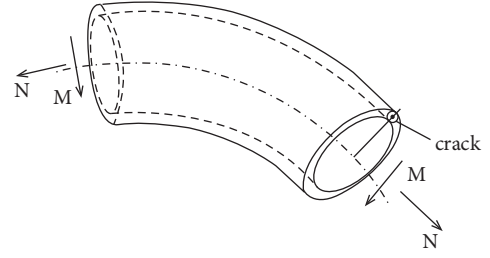


FIGURE 1: Cracked bend under tensile and bending load.

with 6 degrees of freedom, and the elastic modulus $E = 2.1 \times 10^{11}$ and Poisson's ratio $\mu = 0.3$. With the help of KSCON command, the crack tip meshing is established to generate the singular element of stress singularity at the crack tip [12], as shown in Figure 4.

4. BP Neural Network Design

4.1. BP Neural Network Learning Algorithm. BP neural network is the neural network using the error back-propagation algorithm. Its algorithm is as follows: according to the negative gradient direction of the error between the actual output and the expected output of the neural network, the link weight between neurons of each layer is iteratively corrected layer by layer from back to front [13]. Figure 5 shows the three-layer BP neural network model designed in this paper. The learning steps of the neural network backpropagation algorithm are as follows [14]:

- (1) Initialization: set all synaptic weights and thresholds to the minimum random number.
- (2) Provide input training sample set $\{X_i, O_i\}_{i=1}^{160}$, where $X_i = [x_{i1}, x_{i2}, x_{i3}]^T$ and $O_i = [O_i]$ are used as input and output samples, respectively, and select learning step $\eta = 0.6$ to adjust the speed and seismic breaking degree of neural network searching the optimal weight. For $n = 1, 2, \dots, 160$ groups of samples were input circularly.
- (3) For the training sample n , calculate the output of neurons in each hidden layer and output layer:

$$y_j(n) = \phi_j \sum_{i=0}^{mL} w_{ij} y_i(n). \quad (6)$$

- (4) Calculate the error signal $e_j(n) = o_j(n) - y_j(n)$ and the cost function $\varepsilon(n) = (1/2) \sum_{j=1}^{mL} e_j^2(n)$.
- (5) Adjust synaptic weights of output layer and hidden layer:

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji}(n), \quad (7)$$

$$\Delta w_{ji}(n) = \eta \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)}.$$

- (6) Let $n = n + 1$ and return to step (3) until the stop criterion is met.

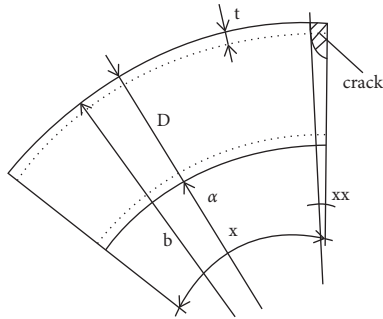


FIGURE 2: Structural dimension of elbow.

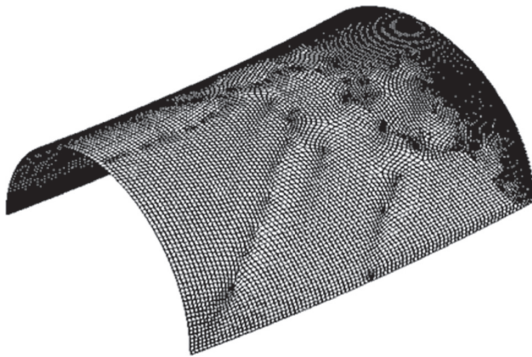


FIGURE 3: Network division diagram of elbow.

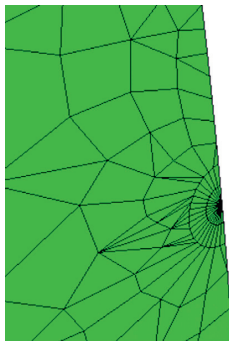


FIGURE 4: Finite element meshing diagram of crack tip with singular element.

4.2. Topology of Prediction Intensity Factor Calculation Network. In this paper, a three-layer BP neural network structure is established, in which there are three nodes in the input layer. The input parameters are the inner diameter a of the elbow, the outer diameter b of the elbow, and the center angle x of the crack. The output is a node, that is, the intensity factor of the cracked elbow. The number of hidden layer nodes affects the accuracy of the network model. At present, there is no theoretical rule for its determination method. The number of hidden layer nodes is related to the amount of input and output information. The network topology is QG3—58—1, as shown in Figure 6. The “tansig” function is used as the activation function, the “purelin” function is used as the activation function in the output layer, and the “trainlm” algorithm is used for training.

5. Experimental Analysis

5.1. Obtaining Training Samples. According to the theory of fracture mechanics, the finite element analysis of the cracked elbow is carried out by ANSYS. The modeling size range is $a = 79\sim 40$ mm, $b = 89\sim 50$ mm; $x = 2.8^\circ$; and $xx = 10^\circ\sim 15^\circ$. In this paper, a singular element is generated at the crack tip of the model, and the stress intensity factor is determined by the three-point displacement extrapolation method at the crack tip.

198 groups of neural network training samples were collected, 159 groups were used for neural network training, and 39 groups were used for neural network verification.

5.2. Result Analysis

5.2.1. Network Training. In this paper, two methods, adding momentum term and adjusting learning rate, are used to improve the learning speed, increase the reliability of the algorithm, and avoid falling into local minimum during neural network training. Adaptively adjusting the learning rate shortens the learning time of neural network. In order to give better play to the performance of the training function, the training samples are quantified to the range of $[-1, 1]$, and the error variation diagram of the network training process shown in Figure 7, as well as the connection weights of each layer and the threshold of each neuron of the network model shown in Table 1, is obtained. Through comprehensive analysis of Figure 7 and Table 1, it can be seen that the network training of this method reaches very high accuracy in step 4500, the mean square deviation of network output is about 2.7902×10^{-7} , and the maximum fitting error is 0.0018, so the network output under the model in this paper is good for target tracking and has high prediction accuracy.

5.2.2. Network Extension Test. The prediction ability of the network to the samples not participating in the training is an effective method to evaluate the reliability of the model [15]. The root mean square error of the intensity factor k was 0.0102, the maximum error was 0.1544, and the average relative deviation of prediction was 0.19%. Figure 8 shows the linear regression analysis diagram of the change rate of the network output change relative to the target value, the linear correlation determination coefficient is $R^2 = 0.9841$, Figure 9 shows the network generalization tracking diagram, and Figure 10 shows the residual diagram of neural network model extension verification. According to the comprehensive analysis of Figures 8–10, the fitting degree between the change of network output and the change rate of target value is good, and the consistency between the predicted value of ANN model and the measured value is good. The error between the predicted value of network and the measured value is mostly distributed in the range of ± 0.15 , and the range of model error is appropriate. Therefore, the neural network prediction model has certain generalization ability and applicability.

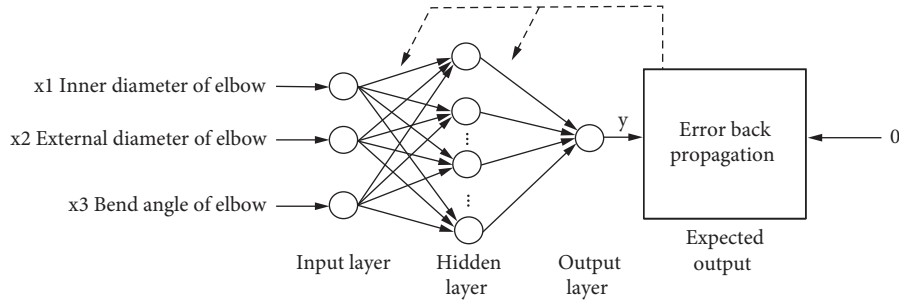


FIGURE 5: BP neural network model.

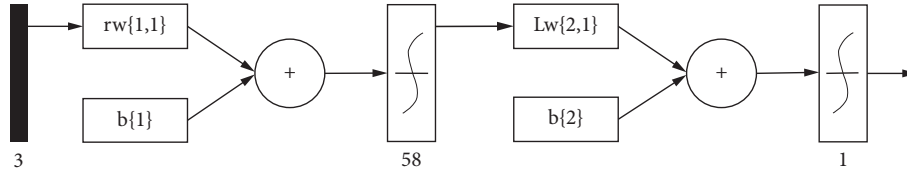


FIGURE 6: Network topology.

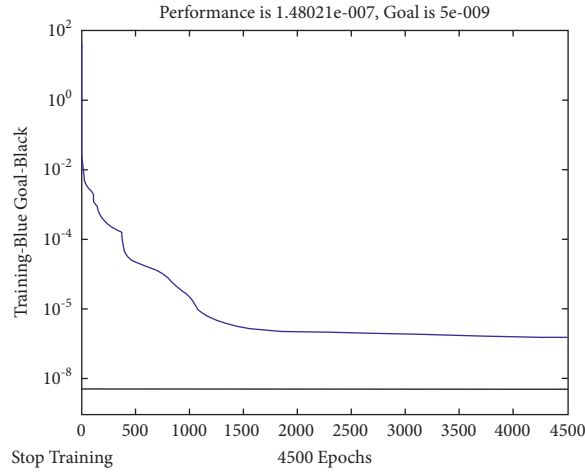


FIGURE 7: Error variation diagram of network training process.

TABLE 1: Network connection weights and thresholds.

Weight	Hidden layer node j					Weight	Hidden layer node j				
	1	2	3	4	5		6	7	8	9	10
Input node i	4.3785	-3.1792	2.4674	-1.1211	-3.8841	Input node i	1.0034	-0.9732	-3.8699	3.6769	-3.2183
	1.2491	-3.3175	2.8006	4.2375	1.6564	Input node i	4.0752	2.5648	3.7926	-0.3202	-3.467
	-2.9388	-2.8731	3.929	3.1867	-3.3968		3.4284	4.6736	-0.0861	3.9681	-2.644
Output node j	0.4383	0.5581	-0.6342	0.5707	-0.9485	Output node j	-0.1852	0.8129	-0.3391	0.839	0.3792
Threshold b	-5.4192	5.2291	-5.0389	4.8488	4.6586	Threshold b	-4.4685	4.2783	4.0882	-3.898	3.7079
Weight	Hidden layer node j					Weight	Hidden layer node j				
	11	12	13	14	15		16	17	18	19	20
Input node i	-3.5057	-0.437	-0.7508	3.7164	2.6689	Input node i	3.09	-2.321	-2.1968	2.9501	5.1046
	-2.5833	-0.5516	-5.2312	2.7578	3.2876	Input node i	-1.8204	4.7142	2.0777	2.2599	-1.1519
	-3.2256	-5.3733	1.1993	2.8197	3.3818		-4.0628	1.3257	-4.4972	3.9443	1.4087
Output node j	0.9053	0.7169	0.644	-0.0452	-0.2266	Output node j	0.9053	0.7169	0.644	-0.0452	-0.2266
Threshold b	3.5177	3.3276	3.1374	-2.9473	-2.7572	Threshold b	3.5177	3.3276	3.1374	-2.9473	-2.7572
Weight	Hidden layer node j					Weight	Hidden layer node j				
	21	22	23	24	25		26	27	28	29	30
Input node i	-3.4534	2.3038	-1.2402	4.7259	3.7538	Input node i	-2.9538	3.4927	3.542	-4.07	3.6032
	2.4541	-4.5069	0.1324	2.2871	-3.8169	Input node i	-4.3738	1.1006	2.3476	-2.42	2.0621
	-3.3792	-1.9362	5.2738	-1.3426	-0.8414		1.2301	-3.9947	3.3632	-2.6357	-3.4832
Output node j	0.4937	-0.8105	-0.8095	0.8194	0.5549	Output node j	-0.7003	-0.7726	-0.5565	0.4276	-0.3155

TABLE 1: Continued.

Weight	Hidden layer node j					Weight	Hidden layer node j				
	1	2	3	4	5		6	7	8	9	10
Threshold b	1.6163	-1.4261	1.236	-1.0458	-0.8557	Threshold b	0.6655	-0.4754	-0.2852	0.0951	0.0951
Weight	Hidden layer node j					Weight	Hidden layer node j				
	31	32	33	34	35		36	37	38	39	40
Input node i	-1.9891	3.3707	-3.0278	-3.6975	-1.3709	Input node i	2.9496	4.0848	-0.5431	-1.8461	-3.8839
Output node j	1.8096	2.1345	-3.1601	3.9596	-5.2084	Output node j	-1.8193	2.9899	4.0422	-3.2275	2.9865
Threshold b	4.705	3.6674	3.196	0.1349	-0.6009	Threshold b	4.1663	-1.9346	3.5684	-3.9425	-2.316
Weight	-0.5323	0.309	-0.4771	-0.1008	0.3014	Weight	-0.6356	-0.0663	0.0656	0.2326	-0.1816
Weight	Hidden layer node j					Weight	Hidden layer node j				
	41	42	43	44	45		46	47	48	49	50
Input node i	-2.1867	-2.3769	2.567	-2.7572	2.9473	Input node i	-3.1374	-3.3276	3.5177	-3.7079	3.898
Output node j	-4.4054	-4.0902	4.1267	-0.8775	3.8594	Output node j	-3.8526	-0.5874	0.9973	-4.6247	0.6137
Threshold b	2.9959	3.5286	-1.8715	2.5035	3.3894	Threshold b	-3.0495	2.2212	4.8828	1.2335	-3.7087
Weight	-0.9926	0.4329	2.9725	-4.7255	-1.7276	Weight	2.286	4.9081	-2.1287	2.5414	-3.9035
Weight	Hidden layer node j					Weight	Hidden layer node j				
	51	52	53	54	55		56	57	58	Output node 1	
Input node i	2.2577	-2.4167	3.017	-5.3484	2.6632	Input node i	-2.6809	-1.3532	-4.0905		
Output node j	3.6244	1.9388	0.1105	-0.6734	4.4392	Output node j	-2.8943	-4.034	-1.7712		
Threshold b	3.3369	4.4462	4.5004	-0.5559	1.6027	Threshold b	3.7154	-3.3562	-3.0819		-0.1685
Weight	0.5577	-0.4144	-0.5542	0.0968	0.728	Weight	-0.0869	0.4252	-0.3411		
Threshold b	4.0882	-4.2783	4.4685	-4.6586	4.8488	Threshold b	-5.0389	-5.2291	-5.4192		

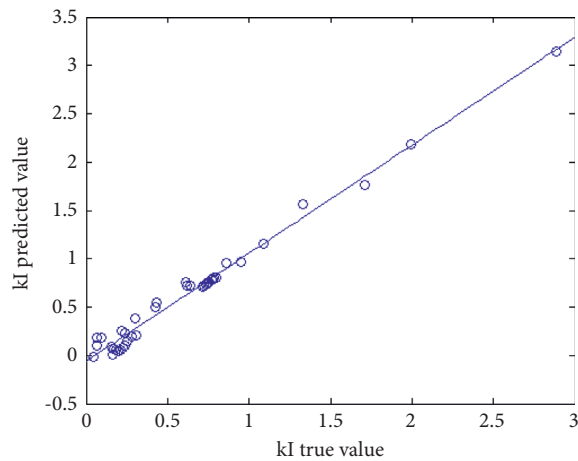


FIGURE 8: Linear regression analysis.

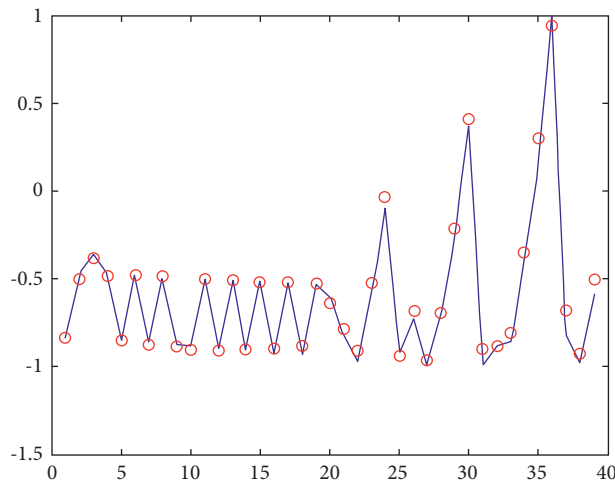


FIGURE 9: Network generalization tracking diagram.



FIGURE 10: Residual diagram of neural network module ductility verification.

6. Conclusion

In order to optimize the accuracy and generalization ability of the existing prediction model of stress intensity factor of circumferential through crack in elbow, a neural network model is introduced to optimize it, and the following conclusions are obtained:

- (1) The training samples are obtained based on the finite element analysis software, and the neural network prediction model of stress intensity factor of cracked elbow is established by using BP neural network. The relative error distribution of fitting and prediction is 0.000027% and 0.19%, which has high fitting and prediction accuracy in a large range.
- (2) Compared with the traditional method, the neural network method is used to establish the calculation model of stress intensity factor of through crack elbow structure, which can effectively avoid the difficulties of other prediction models and the errors caused by human factors.
- (3) BP network is used to replace the traditional calculation method, reduce the calculation difficulty, improve the calculation speed, and speed up the actual engineering structure design and structure analysis, so as to speed up the engineering process.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] K. Ding, L. Wang, W. Wang, K. Wang, B. Ren, and C. Jiang, "Study on the development height of overburden water-flowing fracture zone of the working face," *Geofluids*, vol. 2021, no. 5, pp. 1–10, 2021.
- [2] T. Fett, "Estimation of Stress Intensity Factors for Semi-elliptical Surface cracks," *Engineering Fracture Mechanics*, vol. 66, pp. 349–356, 2000.
- [3] N. -A. Noda, "Stress intensity formulas for three-dimensional cracks in homogeneous and bonded dissimilar materials," *Engineering Fracture Mechanics*, vol. 71, no. 1, pp. 1–15, 2004.
- [4] Li Rong, M. Yang, B. Liang, and N. Nao-Aki, "Calculation method OF stress intensity factor for cracked functionally graded hollow cylinder based ON the ratio OF stresses at crack tip," *Engineering Mechanics*, vol. 37, no. 4, p. 8, 2020.
- [5] B. Qiao, "Computation of crack stress intensity factor based on FEM," *Journal of Xi'an University of Science and Technology*, vol. 30, no. 5, pp. 629–632, 2010.
- [6] Z. Feng, C. Chai, J. Zou, and H. Mou, "Three-dimensional finite element analysis of stress intensity factors of butt-joints with multiple site damage," *Acta Aeronautica et Astronautica Sinica*, vol. 42, no. 5, pp. 237–244, 2021.
- [7] F. Wang, Q. Huang, and Z. Yin, "Analysis of emulational limited element method for the three-dimensional stress intensity factors," *AERONAUTICAL COMPUTING TECHNIQUE*, vol. 36, no. 3, pp. 125–127, 2006.
- [8] M. S. Alam and M. A. Wahab, "Modeling the fatigue crack growth and propagation life of a joint of two elastic materials using interface elements," *International Journal of Pressure Vessels and Piping*, vol. 82, no. 2, pp. 105–113, 2005.
- [9] A. Ajdani, M. R. Ayatollahi, and L. F. M. da Silva, "Mixed mode fracture analysis in a ductile adhesive using semi-circular bend (SCB) specimen," *Theoretical and Applied Fracture Mechanics*, vol. 112, no. 8, Article ID 102927, 2021.
- [10] J. Zhou, W. Zhang, X. Luan, and G. Zhu, "Calculation research of fracture mechanics parameters KJ of cracks in elbows," *Nuclear Techniques*, vol. 36, no. 4, pp. 170–175, 2013.
- [11] X. Lin and A. S. Roderick, "Calculation of stress intensity factors using the 3D finite element method," *China Mechanical Engineering*, vol. 11, no. 9, pp. 39–42, 1998.
- [12] J. Chen, Yi Huang, and G. Liu, "Analysis of finite element model for calculating stress intensity factor based on crack-tip singular element," *SHIPBUILDING OF CHINA*, vol. 51, no. 3, pp. 56–64, 2010.
- [13] Y. Zhang, Z. Kuang, X. Xiao, and B. CHEN, "A direct-weight-determination method for trigonometrically-activated fourier neural networks," *COMPUTER ENGINEERING & SCIENCE*, vol. 31, no. 5, p. 4, 2009.
- [14] H. Yang and W. Gong, "Improvement of back propagation algorithm based on convolution neural network," *Computer Engineering and Design*, vol. 40, no. 1, p. 5, 2019.
- [15] Y. Wang, X. Xie, Y. Guo, and C. Chen, "Cloud resource prediction model based on adaptive neural network," *Science Technology and Engineering*, vol. 21, no. 25, p. 6, 2021.

Research Article

Design of Fault Prediction System for Electromechanical Sensor Equipment Based on Deep Learning

Yongtao Ding , Hua Wu, and Kaixiang Zhou

Department of Mechanical Engineering, City University of Hong Kong, 999077 Hong Kong SAR, China

Correspondence should be addressed to Yongtao Ding; yongt ding2-c@my.cityu.edu.hk

Received 29 November 2021; Accepted 18 January 2022; Published 17 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yongtao Ding 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 increasing complexity, scale, and intelligentization of modern equipment, the maintenance cost of equipment is increasing day by day. Moreover, once an unexpected major failure occurs, it will cause loss and damage to production, economy, and safety. Based on the considerations of system reliability and safety, fault prediction has gradually become a hot topic in the field of reliability. As a new branch of machine learning, deep learning realizes deep abstract feature extraction and expression of complex nonlinear relations by stacking deep neural networks and makes its methods solve bad problems in many traditional machine learning fields. The improvement and excellent results have been achieved. This article first introduces the model structure and working principle of the classic deep learning model noise reduction autoencoder and combines the feature extraction results of the experimental data of electromechanical sensor equipment and the model characteristics to analyze that this type of model can be trained using only normal samples. Under the restriction, the reason for abnormal features can also be correctly filtered. Then, in order to suppress the overfitting of training, the dropout layer is used in this article. The dropout layer will make the hidden layer nodes to be dropped with probability p according to the set probability. Because the lost nodes are random, each training is equivalent to training a new model. It achieves an effect similar to independently training the model and then superimposing it. Experiments prove that the dropout layer is very effective in solving overfitting. Finally, this paper conducts experimental verification on the improved algorithm. The results show that the improved model has a certain improvement in accuracy under the limited training algebra. Under the same training parameters, the accuracy is increased by approximately 2.44%, and the improved model has a better training effect and can be used for electromechanical effective prediction of sensor equipment failure.

1. Introduction

The structure of large-scale mechanical equipment such as high-speed EMU trains, steam turbines, and aerospace engine units is becoming more and more complex, and the probability of failure is gradually increasing. Fault diagnosis helps the system operate safely and efficiently. However, the traditional planned maintenance needs to stop production and cause high maintenance costs, and no matter how the maintenance interval is set, there is always the problem of insufficient maintenance or overmaintenance. Therefore, in consideration of factors such as reliability, safety, and economy of complex systems, equipment predictive maintenance provides maintenance engineers with a decision-making

basis through early fault diagnosis and remaining life prediction so that the transition from planned maintenance to condition-based maintenance can be realized. United Airlines' predictive maintenance strategy has reduced the maintenance cost of Douglas DC-8 aircraft by 30%, and the fatality rate per 100 million passenger-kilometers has dropped from 0.6 to about 0.2. After adopting the predictive maintenance strategy, the United States' fifth-generation fighter aircraft not only reduced maintenance costs but also increased the overall reliability by more than 50%. However, the accuracy of fault diagnosis and remaining life prediction depends on the effectiveness of fault feature extraction methods. As an advanced feature extraction tool, deep learning has been widely used in natural language processing, text

processing, and other fields. In recent years, it has also attracted the attention of experts and scholars in the field of fault diagnosis [1–6].

Starting from the original monitoring data of electromechanical sensor equipment, this paper studies the fault diagnosis method of electromechanical sensor equipment based on deep learning. First, build a deep neural network prediction framework based on understanding the principles of deep learning; then start from the electromechanical sensor monitoring information; study equipment fault diagnosis methods for electromechanical sensor monitoring information; and finally, on this basis, further study more complex multisensors. The monitoring information fault diagnosis lays the foundation for the subsequent research on the remaining life prediction of equipment based on deep learning.

2. Related Work

With the development of equipment in the direction of complexity, intelligence, and integration, its working environment and operating conditions have become more and more complex, and the data collected by the monitoring system have become larger and larger, based on the current large amount of electromechanical data, the fault diagnosis of electromechanical sensor equipment has become the research focus of machine learning. It has become a hot issue in the current PHM field. Drawing on the successful experience of deep learning, some scholars have already applied deep learning in the field of fault diagnosis and have achieved certain results. Since Tamilselvan first applied DBN to aircraft engine fault diagnosis in 2013, more and more scholars at home and abroad have paid attention to this field and achieved many research results. Tran et al. merged the DBN and TKEO algorithms and applied them to the valve fault diagnosis of reciprocating compressors. The author used GRBM instead of RBM to get rid of the shortcomings of traditional RBM that can only input binary fault signals and use the DBN model as a classifier for electromechanical sensors and a traditional Residual Neural Network, a higher accuracy rate has been achieved. Jia et al. applied deep neural networks to the feature extraction and diagnosis of bearing faults. They used the feature extraction capabilities of deep neural networks to extract and identify bearing fault features directly from frequency domain signals and achieved good results. Lu et al. specially studied the fault feature extraction ability of deep neural networks in the frequency domain signals of rotating bearings and gave the visualization results of the fault features. By summarizing the above-mentioned literature, although deep learning has been successfully applied in the field of failure, the research in this field is still immature, and there are still many problems to be solved. The main problems are summarized as follows: (1) only use deep learning as a classifier, traditional signal processing methods are still used to extract fault features, and the ability of deep learning to mine fault features is not fully utilized; (2) when using deep learning algorithms to

extract fault features and identify the health status of fault signals, most of the signal frequencies are used, and when the signal does not have periodicity, this method will fail and is not general; (3) the comparative study of deep learning and other shallow models is not sufficient; and (4) the role of deep learning in the fault diagnosis of electromechanical sensor monitoring information application remains to be studied [7–15].

3. Related Theoretical Methods

3.1. Data-Driven Remaining Life Prediction Framework. A relatively complete data-driven RUL prediction framework suitable for electromechanical sensor equipment systems often covers two stages of online application and offline modeling. The online and offline parts complement and support each other. In the offline phase, make full use of equipment system test data, historical monitoring data, physical and mathematical models of subsystems or components, and so on to establish operating condition recognition, data preprocessing, feature engineering, degradation status recognition, and RUL for online practical applications. In the online stage, after necessary data preprocessing, feature engineering and other operations are performed on the data collected in real-time monitoring, based on the actual operating environment of the service equipment system (such as external environment, load conditions, etc.), with the help of intelligent algorithmic prediction system RUL. In this framework, training and learning in the offline phase and actual prediction in the online phase are fully combined, and the model can be adjusted in an offline or online dynamic manner [16–20], as shown in Figure 1.

Data-driven RUL prediction methods are currently mainly based on machine learning, mathematical statistics, and deep learning. Different prediction methods have different adaptabilities. Corresponding to condition monitoring data, there are usually two RUL prediction strategies. One is to map m -dimensional features to a single-dimensional HI value curve and use curve fitting, extrapolation, and other methods to predict, which is called RUL indirect prediction strategy. This article uses direct pattern matching on the original multidimensional data and extracts the m -dimensional characterization of degraded fault features to predict the remaining life of the equipment system, which is an RUL direct prediction strategy.

3.2. Deep Neural Network. A deep neural network can be regarded as a multihidden neural network formed by stacking multiple AEs. It uses bottom-up unsupervised learning to extract features layer by layer and uses supervised learning methods to fine-tune the parameters of the entire network. The most essential characteristics of a certain state of the object are extracted from the original data, and the deep neural network structure is shown in Figure 2 [21].

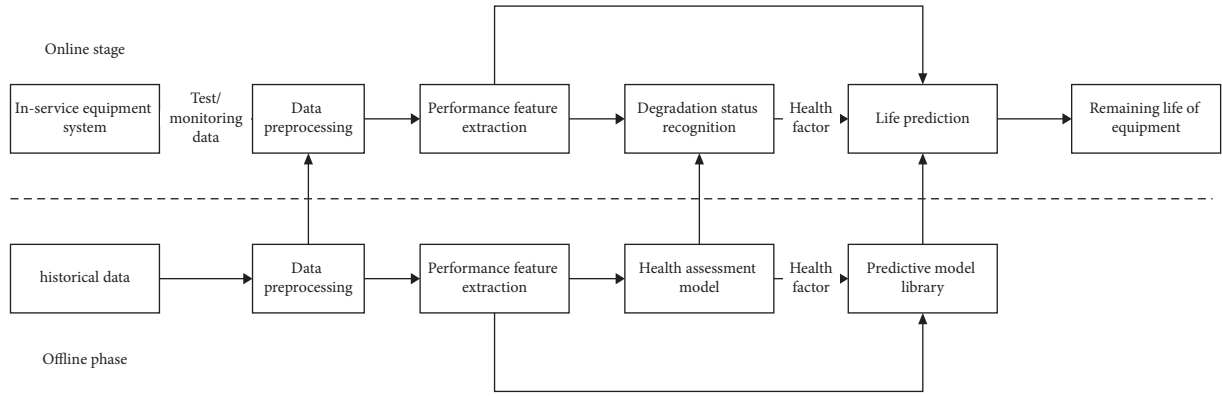


FIGURE 1: General framework for data-driven remaining life prediction.

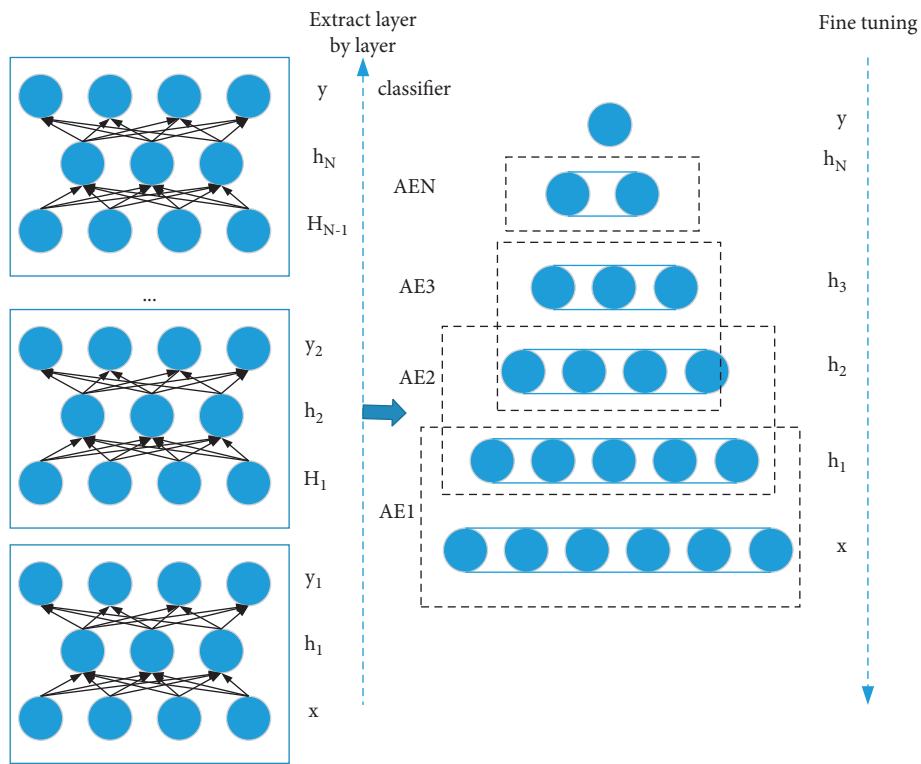


FIGURE 2: Deep neural network model composed of multiple AEs.

3.2.1. Autoencoder. The autoencoder is a three-layer forward artificial neural network, as shown in Figure 3. Its output is equal to the input, and it is mainly composed of an input layer, a hidden layer, and an output layer. The autoencoder is mainly composed of two parts: encoder and decoder. The input data are transformed into features through the encoder network, and the input data are encoded from high- to low-dimensional space data, and the input data are obtained in low-dimensional space. Then, the low-dimensional space data are mapped to the high-dimensional space through the decoder network to realize the reconstruction of the input to the output, and the obtained low-dimensional space data can be used as the feature representation of the input high-dimensional space data [22].

Given an unlabeled data set containing P observation variables and M samples $\{x_{pm}\}$, ($p = 1, 2, \dots, p; m = 1, 2, \dots,$

M), the coding network passes The activation function f_θ encodes the sample x_m into the activation value h_m of the hidden layer, and the coding process is shown in the following equation:

$$h_m = f_\theta(x_m) = \sigma(Wx_m + b). \quad (1)$$

f_θ is the coding function, the function s usually takes the Sigmoid function as the activation function of the coding network, W is the network weight matrix between the input layer and the hidden layer, b is the bias vector generated by the coding network, $\theta = \{W, b\}$ is the connection weight and bias parameter between the input layer and the hidden layer. The general form of the Sigmoid function is shown in the following formula:

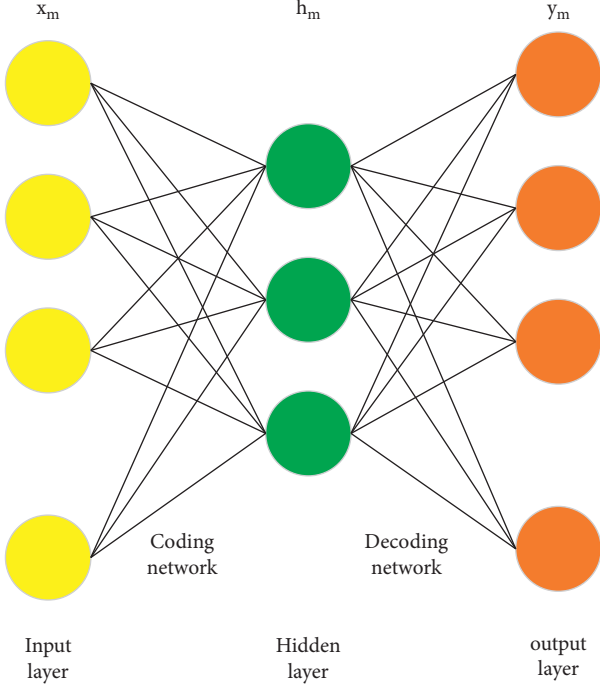


FIGURE 3: Autoencoder model.

$$\sigma x = \frac{1}{1 + e^{-x}}. \quad (2)$$

Similarly, for the decoding network, the encoding vector h obtained by the encoding network is reconstructed by the decoding network to obtain m_y , that is equal to the input, that is, m_y and the input m_x are equal. The decoding process is shown in the following equation:

$$y_m = g_{\theta^T}(h_m) = \sigma(W^T h_m + d). \quad (3)$$

g_{θ^T} is the decoding function, s is the activation function of the encoding process, W^T is the network weight matrix from the hidden layer to the output layer, and d is the bias vector generated during the encoding process. The essence of the process of training AE is to train and optimize the network parameters θ and θ^T . In order to make the output y_m as close as possible to the input x_m , by minimizing the reconstruction error $J(\theta, \theta^T)(x, y, W, b)$ to characterize the closeness between the input and the output, as shown in the following equation:

$$J_{(\theta, \theta^T)}(x, y, W, b) = \frac{1}{m} \|y - x\|^2. \quad (4)$$

In each training process, the gradient descent method is used to update the network training parameters θ and θ^T of AE. The entire parameter update process is as follows:

$$W_l = W_l - \alpha \frac{\partial}{\partial W_l} J_{(\theta, \theta^T)}(x, y, W, b), \quad l = 1, 2, \quad (5)$$

$$b_l = b_l - \alpha \frac{\partial}{\partial b_l} J_{(\theta, \theta^T)}(x, y, W, b), \quad l = 1, 2.$$

α is the learning rate and $\alpha \frac{\partial}{\partial W_l} J_{(\theta, \theta^T)}(x, y, W, b)$ and $\alpha \frac{\partial}{\partial b_l} J_{(\theta, \theta^T)}(x, y, W, b)$ are calculated using backpropagation algorithm.

3.2.2. Deep Neural Network Training. First, the DNN network is pretrained for the failure prediction of electromechanical sensor equipment through the training algorithm. Train the first autoencoder AE_1 by giving an unlabeled input data set x as the input of the coding network and get the coding vector h_1 . Set x as the output of AE_1 to obtain the training parameter θ_1 ; and then use the hidden layer feature h_1 of the first autoencoder as the input of the second autoencoder AE_2 ; and train to obtain the network parameter θ_2 of AE_2 and h_2 as the hidden layer of AE_2 . The data can be viewed as a feature representation of AE_2 . Repeat this process to obtain the hidden layer feature h_N of the N -th autoencoder AE_N and the corresponding network training parameter θ_N [23, 24].

Second, add a classifier to the top of the DNN network. The pretraining process of the DNN is completed through the layer-by-layer unsupervised training method, which realizes the layer-by-layer extraction of feature information. However, the DNN at this time does not have a classification function. In order to achieve the output classification function, a classifier needs to be added to the top layer of the DNN. This paper uses the softmax classifier as the output layer of DNN. Assuming that the training data set is $\{x_m\} (m=1, 2, \dots, M)$, the label is $u_m \in \{1, 2, \dots, k\}$; each probability $p(u=i|x)$ of a type $i (i=1, 2, \dots, k)$ can be calculated by the following hypothesis function:

$$h_{\theta} x_m \begin{bmatrix} pu_m = 1 | x_m; \theta \\ pu_m = 2 | x_m; \theta \\ \vdots \\ pu_m = k | x_m; \theta \end{bmatrix} = \frac{1}{\sum_{i=1}^k e^{\theta_i^T x_m}} \begin{bmatrix} e^{\theta_1^T x_m} \\ e^{\theta_2^T x_m} \\ \vdots \\ e^{\theta_k^T x_m} \end{bmatrix}. \quad (6)$$

θ is the model parameter of softmax, which is similar to the AE model. In order to ensure the performance of the classifier, the training parameters of the model are optimized by minimizing the loss function J_{θ} . The top-level network parameter θ_{N+1} can be obtained by minimizing $J_{\theta} x_m$ as follows:

$$J_{\theta} x_m = -\frac{1}{M} \left[\sum_{m=1}^M \sum_{m=1}^M 1\{u_m = i\} \log \frac{e^{\theta_i^T x_m}}{\sum_{i=1}^k e^{\theta_i^T x_m}} \right]. \quad (7)$$

Finally, fine-tune the training parameters of the entire DNN. In order to ensure the accuracy of feature extraction and the classification effect of the output layer, the backpropagation algorithm is used to supervise the fine-tuning of the entire DNN training parameters through a limited number of sample labels, and the fine-tuning process is completed by minimizing the reconstruction error $E(\theta)$. The parameter update process is as follows:

$$E(\theta) = \frac{1}{M} \sum J_{\theta}(Y_m, u_m; \theta), \quad (8)$$

$$\theta = \theta - \alpha \frac{\partial E(\theta)}{\partial (\theta)}.$$

Y_m is the true output value, θ is the parameter set obtained from the entire network training process $\theta = \{\theta_1, \theta_2, \dots, \theta_N, \theta_{N+1}\}$, the back-propagation algorithm is mainly used to continuously modify the network parameters θ , and α is the learning rate. The fine-tuning process uses labeled data to improve the accuracy of DNN training.

4. Based on the Improved Residual Neural Network Electromechanical Sensor Equipment Fault Diagnosis Model System Design

4.1. Random Weight Average Method SWA

4.1.1. Learning Rate. During SGD training, there are two strategies for setting the learning rate: fixed learning rate and periodic learning rate. Under a certain learning rate strategy of SWA, the average of the weight data in the SDP training process is used as the weight algorithm. When the environment of the learning rate changes, assuming that the learning rate decreases linearly from α_1 to α_2 in a cycle, the change formula of the learning rate is

$$\alpha(i) = 1 - t\alpha_1 + t(i)\alpha_2 = \frac{1}{c} (\text{mod}(i-1, c) + 1). \quad (9)$$

4.1.2. Learning Rate Error Curve. The performance of the learning rate and error rate during the training of the SGD algorithm is shown in Figure 4.

In Figures 1–4, the picture shows the change curve of the learning rate, and the picture below shows the training error corresponding to the learning rate. The small circle on the learning rate curve indicates the minimum learning rate and its corresponding test error. Figure 4 shows the change in the learning rate and the corresponding error rate during a batch of training. It can be found that the learning rate does not change continuously. At the end of a cycle, the learning rate jumps directly from the minimum to the maximum. Basically, the learning rate $\alpha_1 \geq \alpha_2$, and the length c of a cycle is a hyper-parameter. In order to improve training, it is also necessary to consider the case where the learning rate remains unchanged.

4.2. Dropout Regularization. If the training sample is insufficient, it is easy to produce overfitting phenomenon. The phenomenon of overfitting will reduce the effect of actual prediction. The failure prediction model of electromechanical sensor equipment has a smaller loss on the training set and a higher accuracy rate, while the opposite is true on the test machine. Dropout not only can effectively reduce the overfitting phenomenon but also can reduce the fault prediction error rate of electromechanical sensor equipment. Its essence is that during training, the neuron becomes 0 with probability p and does not participate in the process of forward propagation. This brings about a reduction in the dependence of the electromechanical sensor equipment failure prediction model on local features, so the electromechanical sensor equipment failure prediction model has better generalization capabilities. The structure of dropout is shown in Figure 5.

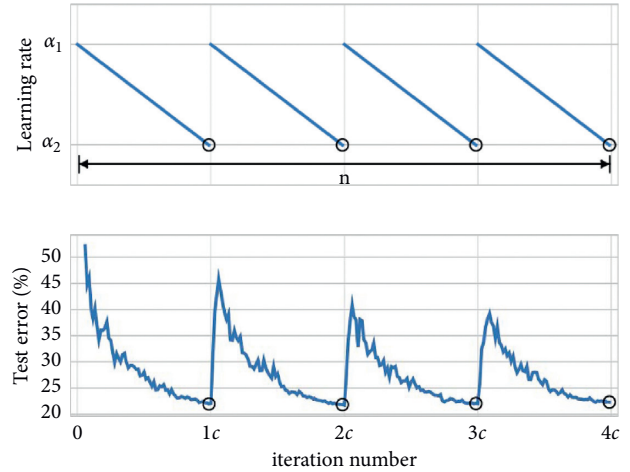


FIGURE 4: Periodic change learning rate and its training error rate.

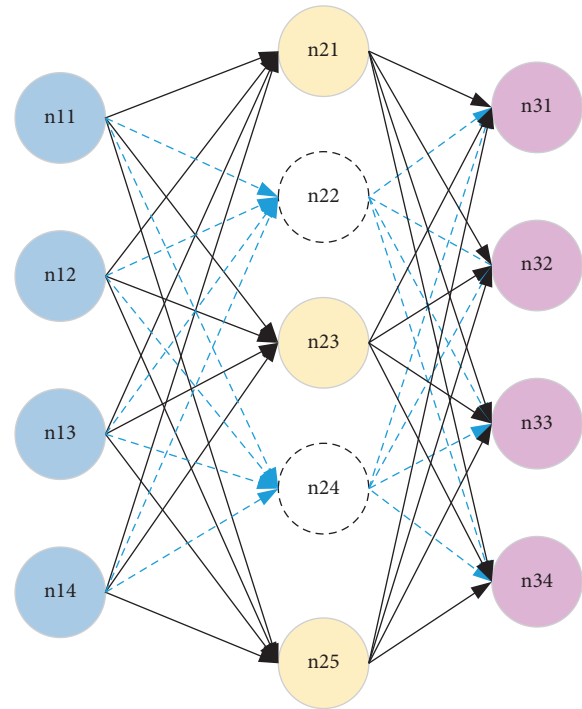


FIGURE 5: Network structure using dropout.

4.2.1. Workflow of Dropout

- (1) Set the node hidden probability $0 < p < 1$, and the neurons in the hidden layer will be shielded with probability p
- (2) The neurons in the hidden layer are shielded with probability p , and the shielded neurons no longer participate in the training process
- (3) Neurons that have not been deleted proceed forward and backward in the normal way and update normally (w, b)
- (4) Restore the shielded neurons and then perform steps 1 and 2

From the above analysis, we can know that the idea of dropout is that certain neurons do not participate in training at certain times.

4.2.2. How to Implement Dropout. During the training phase, each neuron needs to be shut down probabilistically. The model diagram is shown in Figure 6. The corresponding calculation process is as follows.

When there is no dropout layer, the calculation becomes

$$z_i^{(l+1)} = w_i^{(l+1)} y^j + b_i^{(l+1)}. \quad (10)$$

The corresponding network calculation formula without dropout is as follows:

$$r_i^{(l)} \sim \text{Bernoulli}(p), \quad (11)$$

$$y^j{}^{(l)} = r^{(l)} \times y^{(l)}, \quad (12)$$

$$z_i^{(l+1)} = w_i^{(l+1)} y^j{}^{(l)} + b_i^{(l+1)}, \quad (13)$$

$$y_i^{(l+1)} = f(z_i^{(l+1)}). \quad (14)$$

Equations (12)–(14) use Bernoulli functions. The function of the Bernoulli function is to generate a random vector composed of 0 and 1 according to the probability. In addition, f represents the activation function; y represents the output; z represents the value of the neuron after the summation, which becomes the output after the activation function.

4.2.3. Effectiveness Analysis of Dropout. For model optimization, there is a classic method. For a specific electromechanical sensor device to predict the scene, train multiple models at the same time, as shown in Figure 7. This is a three-category model that predicts three categories of $a/b/c$, among which

$$\alpha = \frac{\alpha_1 + \alpha_2}{2}, \quad (15)$$

$$b = \frac{b_1 + b_2}{2},$$

$$c = \frac{c_1 + c_2}{2}. \quad (16)$$

Figure 7 shows a structure in which two models are averaged. Through the superposition of multiple networks, the generalization ability of the predictive model of electromechanical sensor equipment can be improved. And dropout achieves the effect of multiple networks superimposed to a certain extent. Because after regularization, the network will randomly lose some neurons each time, the model structure after discarding is different from that before discarding, which is equivalent to a new network. In this way, there will be many different networks, and then the final prediction results will be obtained by averaging or voting on these models. Dropout replaces the multimodel solution to a certain extent. From the above analysis, we know that some neurons will be randomly closed every time forward and backward. In essence, the neurons can be considered as a new network after they are

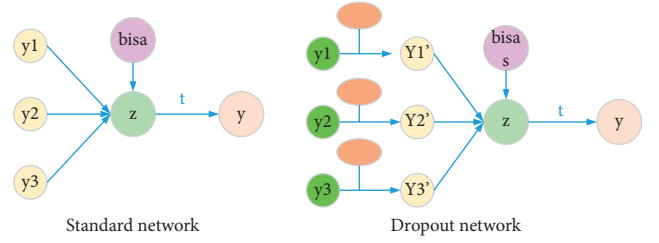


FIGURE 6: Comparison of the ordinary network and dropout layer.

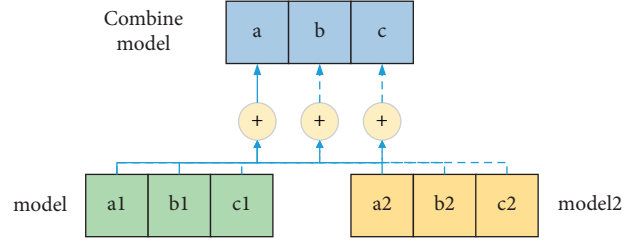


FIGURE 7: The fusion of the two models.

closed. In this way, multiple networks are generated during the training process, and the final prediction result is the average effect of multiple networks. After taking the average, the risk of overfitting of the model is effectively suppressed. With the introduction of dropout, two neurons may not appear in the network at the same time. In this way, the weight update does not depend on the joint action of all hidden nodes. In this way, the problem of excessive reliance on special features in the training process is prevented. Make the neural network train some more common features for classification. Increase the robustness of the network. It can also be said that the introduction of dropout reduces the sensitivity of special features during the training process so that the generalization ability of the system is significantly improved. The dropout layer prevents some nodes from participating in training so that the actual training time of the node is less than the set algebra, which also reduces the risk of overfitting in the prediction process of electromechanical sensor equipment.

5. Fault Prediction and Analysis of Electromechanical Sensors Based on Improved Residual Neural Network

5.1. Operation of Electromechanical Sensor Fault Prediction System Operation. For the residual neural network, building the residual block is the most fundamental task of building a residual neural network. A good residual block design is very important for neural networks. A typical residual block contains two convolutional layers, and the two convolutional layers are not directly connected. In order to improve the recognition effect of the algorithm, we introduced the BN layer and the activation function, and the combination between them constitutes a complete residual layer. There are multiple combinations of residual blocks, and the residual structure selected in this paper is shown in Figure 8 [25, 26].

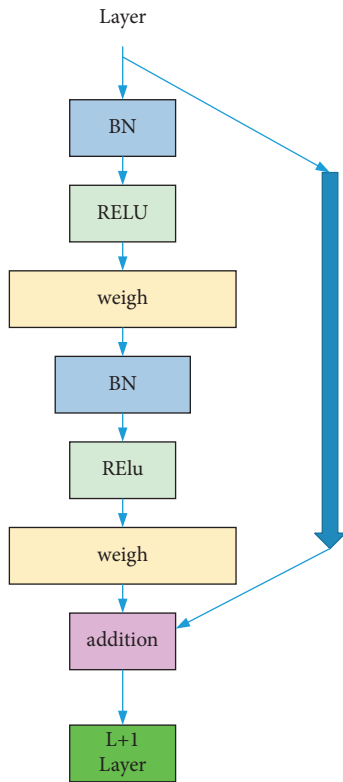


FIGURE 8: Structure realization of the residual block.

This paper analyzes the predictions of electromechanical sensor equipment for improved networks, optimized by swa algorithm, dropout regularization, and both algorithms. After the dropout optimization of the ResNet50 network structure, its network structure becomes Figure 9. In ResNet50, a total of five dropout structures are added.

5.2. Result Analysis. This paper randomly selects 30 fault features; the training algebra is set to 30 generations; and the trained electromechanical sensor equipment prediction model is ResNet50 and its improved model. After comparison of experiments, it is found that the four kinds of electromechanical sensor equipment prediction models trained have overfitting, but the difference is that their overfitting degrees are inconsistent. The following are the experimental results: the training results of the original ResNet50 are shown in Figure 10.

The training results are as follows: The training accuracy rate is 99.84%, and the loss is 0.03. The training accuracy rate is 91.69%, and the loss is 0.49.

After adding dropout regularization, the training effect is shown in Figure 11.

The training results are as follows: The training accuracy rate is 99.93%, and the loss is 0.04. The training accuracy rate is 92.75%, and the loss is 0.67.

After improving the model through the swa algorithm, the training effect diagram is shown in Figure 12.

The training results are as follows: The training accuracy rate is 99.82%; loss is 0.02.

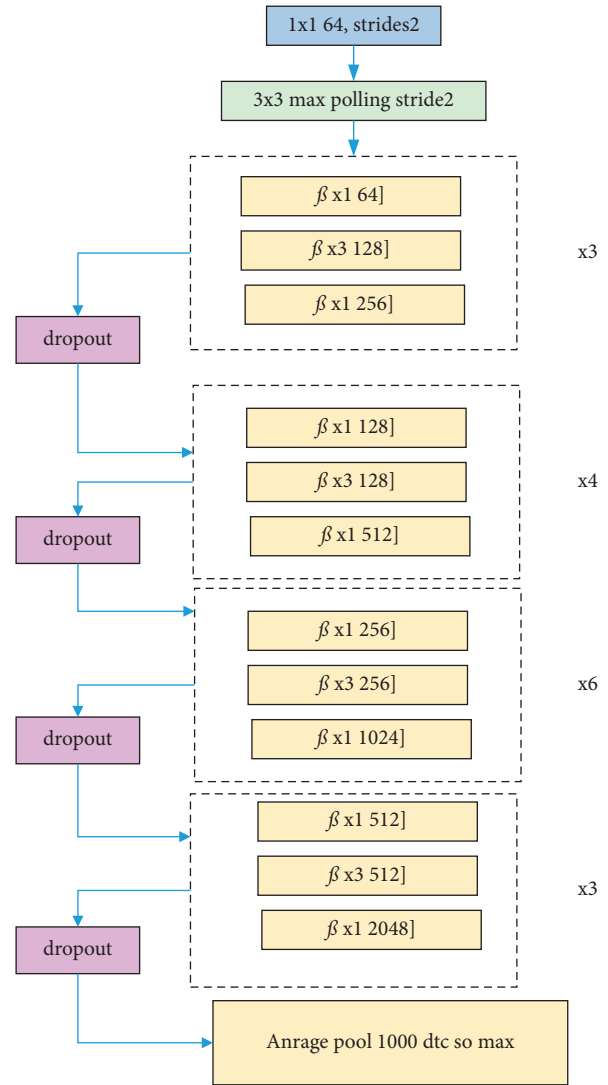


FIGURE 9: ResNet50 network structure after adding dropout.

The training accuracy rate is 95.19%; loss is 0.54. After adding dropout regularization and swa algorithm improvement, the training effect diagram is shown in Figure 13.

The training results are as follows: The training accuracy rate is 99.51%, and the loss is 0.3. The training accuracy rate is 95.61%, and the loss is 0.44.

Through the above data verification, the following conclusions can be drawn:

- (1) The four electromechanical sensor equipment prediction models all have an overfitting phenomenon, that is, the test effect of the test set is significantly lower than the training effect of the training set.
- (2) To the network that joins the dropout, the training process fluctuates greatly, which is in line with the expected result. Because when the network turns off certain weights, it is equivalent to a new network, and training on this basis may fluctuate.

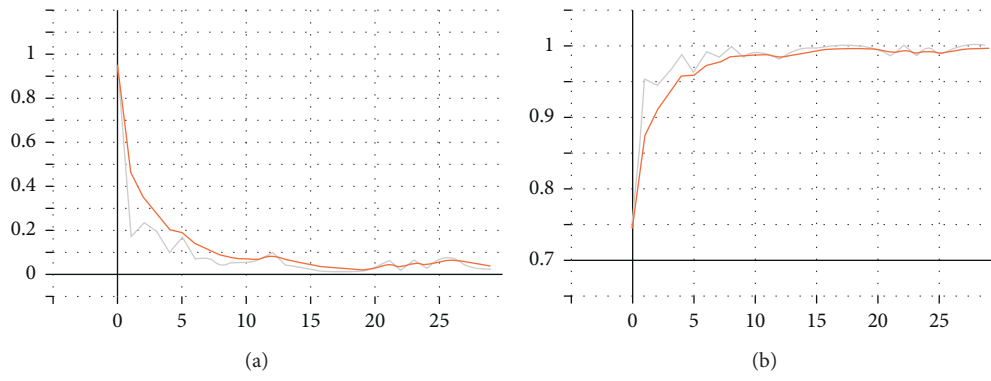


FIGURE 10: ResNet training effect, left loss, and right Acc before the improvement.

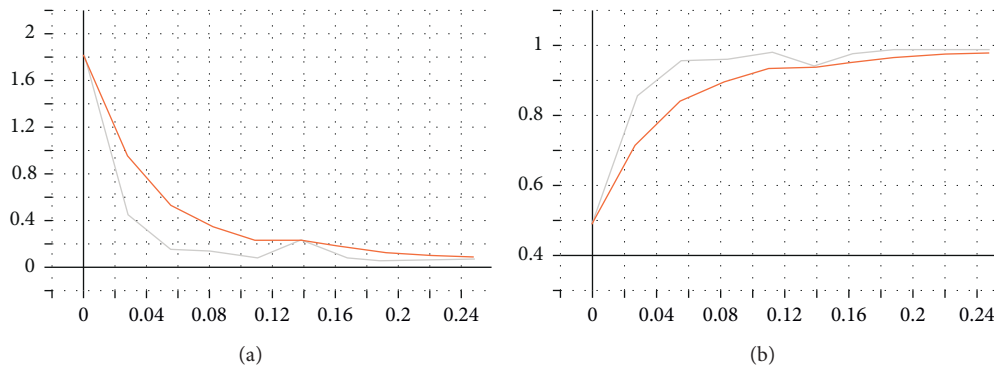


FIGURE 11: ResNet training effect, left loss, and right Acc after adding dropout regularization.

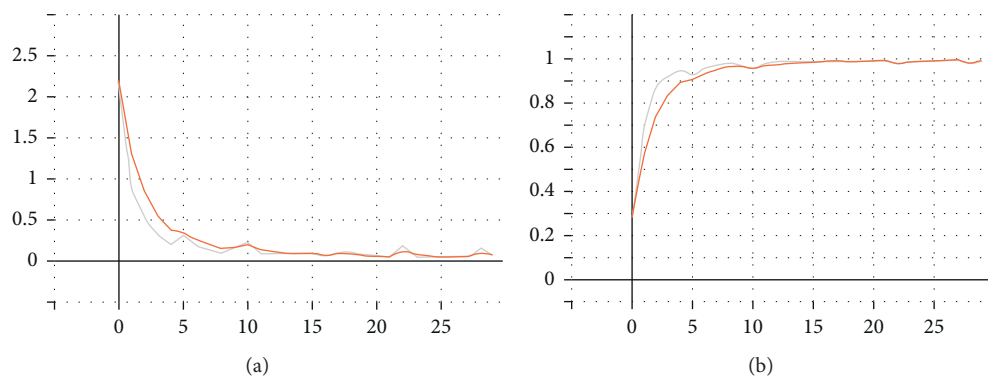


FIGURE 12: swa improved ResNet50 training results.

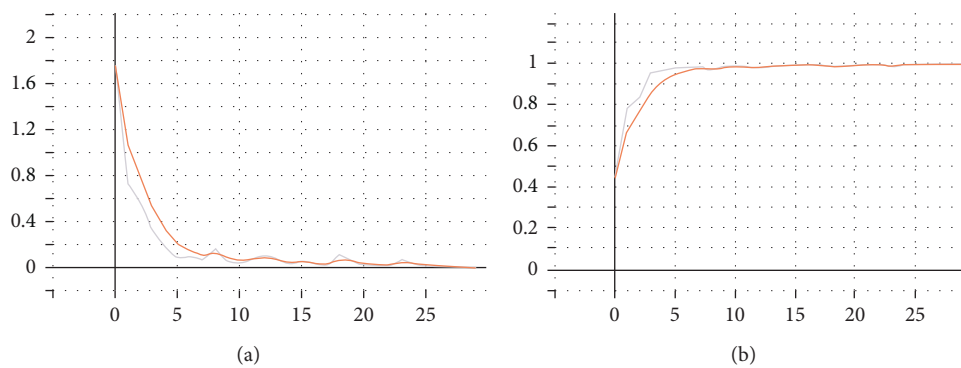


FIGURE 13: Improved swa and dropout, ResNet50 training results.

- (3) Networks that use swa algorithm improvement and simultaneous swa algorithm and dropout regularization algorithm improvement have the best test results, but the latter is slightly better. This shows that the random weight averaging algorithm and dropout have the effect of suppressing overfitting.

6. Conclusion

This article takes electromechanical sensor equipment as the research object. Aiming at the situation of strong human intervention and insufficient recognition ability in the fault diagnosis of electromechanical sensor equipment, the study of deep learning in the fault diagnosis of electromechanical sensor equipment is carried out. The work of this article is summarized as follows:

The advanced convolutional neural network model is used to maximize the use of the fault recognition ability of the deep learning network for fault diagnosis, reducing manual intervention models.

Aiming at the phenomenon that the electromechanical sensor equipment predicts that the training process is prone to overfitting, two improved methods are adopted. Based on the random weight averaging algorithm and dropout regularization, the model is improved from two aspects: backpropagation and model structure. Experiments show that when the backpropagation learning rate is fixed or periodically changes, the SWA algorithm can effectively improve the training effect. Dropout regularization is to randomly discard some neurons so that the network will not rely too much on certain local features, thereby improving the generalization ability of the model to suppress overfitting, and the effect of electromechanical sensor equipment failure prediction is better.

Data Availability

The data set can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] A. El Mejdoubi, H. Chaoui, H. Gualous, P. Van Den Bossche, N. Omar, and J. Van Mierlo, "Lithium-ion batteries health prognosis considering aging conditions," *IEEE Transactions on Power Electronics*, vol. 34, no. 7, pp. 6834–6844, 2019.
- [2] Z. Che, S. Purushotham, K. Cho, D. Sontag, and Y. Liu, "Recurrent neural networks for multivariate time series with missing values," *Scientific Reports*, vol. 8, no. 1, p. 6085, 2018.
- [3] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067–7075, 2016.
- [4] O. Janssens, V. Slavkovikj, B. Vervisch et al., "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, 2016.
- [5] P. Baldi, P. Sadowski, and D. Whiteson, "Searching for exotic particles in high-energy physics with deep learning," *Nature Communications*, vol. 5, Article ID 4308, 2014.
- [6] P. Izmailov, D. Podoprikin, T. Garipov, D. Vetrov, and A. G. Wilson, "Averaging weights leads to wider optima and better generalization," vol. 1803, pp. 1–12, 2018, <https://arxiv.org/abs/1803.05407>, Article ID 5407.
- [7] C. Li, R.-V. Sánchez, G. Zurita, M. Cerrada, and D. Cabrera, "Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning," *Sensors*, vol. 16, no. 6, p. 895, 2016.
- [8] S. Shili, A. Hijazi, A. Sari, X. Lin-Shi, and P. Venet, "Balancing circuit new control for supercapacitor storage system lifetime maximization," *IEEE Transactions on Power Electronics*, vol. 32, no. 6, pp. 4939–4948, 2017.
- [9] S. Subramanian, M. A. Johnny, M. Malamal Neelanchery, and S. Ansari, "Self-Discharge and voltage recovery in graphene supercapacitors," *IEEE Transactions on Power Electronics*, vol. 33, no. 12, pp. 10410–10418, 2018.
- [10] A. El Mejdoubi, H. Chaoui, H. Gualous, and J. Sabor, "Online parameter identification for supercapacitor state-of-health diagnosis for vehicular applications," *IEEE Transactions on Power Electronics*, vol. 32, no. 12, pp. 9355–9363, 2017.
- [11] F. Naseri, E. Farjah, M. Allahbakhshi, and Z. Kazemi, "Online condition monitoring and fault detection of large supercapacitor banks in electric vehicle applications," *IET Electrical Systems in Transportation*, vol. 7, no. 4, pp. 318–326, 2017.
- [12] H. Yang, "Estimation of Supercapacitor charge capacity bounds considering charge redistribution," *IEEE Transactions on Power Electronics*, vol. 33, no. 8, pp. 6980–6993, 2018.
- [13] W. Houlian and Z. Gongbo, "State of charge prediction of supercapacitors via combination of kalman filtering and backpropagation neural network," *IET Electric Power Applications*, vol. 12, no. 4, pp. 588–594, 2018.
- [14] B. Dolenc, P. Bošković, M. Stepančič, A. Pohoranta, and D. Juričič, "State of health estimation and remaining useful life prediction of solid oxide fuel cell stack," *Energy Conversion and Management*, vol. 148, pp. 993–1002, 2017.
- [15] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Prognostics of proton exchange membrane fuel cell stack in a particle filtering framework including characterization disturbances and voltage recovery," in *Proceedings of the IEEE Conference on Prognostics and Health Management*, pp. 1–6, Cheney, WA, USA, June 2014.
- [16] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Pera, and N. Zerhouni, "Joint particle filters prognostics for proton exchange membrane fuel cell power prediction at constant current solicitation," *IEEE Transactions on Reliability*, vol. 65, no. 1, pp. 336–349, 2016.
- [17] M. Ahwiadi and W. Wang, "An enhanced mutated particle filter technique for system state estimation and battery life prediction," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 3, pp. 923–935, 2019.
- [18] F. Cadini, C. Sbarufatti, F. Cancelliere, and M. Giglio, "State-of-life prognosis and diagnosis of lithium-ion batteries by data-driven particle filters," *Applied Energy*, vol. 235, pp. 661–672, 2019.
- [19] D. Andre, C. Appel, T. Soczka-Guth, and D. U. Sauer, "Advanced mathematical methods of SOC and SOH estimation for lithium-ion batteries," *Journal of Power Sources*, vol. 224, pp. 20–27, 2013.
- [20] K. Cho, B. V. Merriënboer, C. Gulcehre et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *Computer Science*, 2014.

- [21] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, pp. 386–405, 2018.
- [22] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, and A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of li-ion batteries," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 8, pp. 6730–6739, 2018.
- [23] S. K. Ibrahim, A. Ahmed, M. A. E. Zeidan, and I. E. Ziedan, "Machine learning methods for spacecraft telemetry mining," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 4, pp. 1816–1827, 2019.
- [24] T. Ergen and S. S. Kozat, "Efficient online learning algorithms based on LSTM neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 8, pp. 3772–3783, 2018.
- [25] M. Elhoseny, A. Tharwat, A. Farouk, and A. E. Hassanien, "K-coverage model based on genetic algorithm to extend WSN lifetime," *IEEE sensors letters*, vol. 1, no. 4, pp. 1–4, 2017.
- [26] L. Fabbri and M. H. Wood, "Accident damage analysis module (ADAM): novel European commission tool for consequence assessment-scientific evaluation of performance," *Process Safety and Environmental Protection*, vol. 129, pp. 249–263, 2019.

Research Article

Dynamic Evolution Analysis of Desertification Images Based on BP Neural Network

Guanyao Lu , Dan Xu, and Yue Meng 

School of Environmental and Chemical Engineering, Foshan University, Foshan 528000, Guangdong, China

Correspondence should be addressed to Guanyao Lu; guanyaolu@fosu.edu.cn

Received 17 January 2022; Revised 28 January 2022; Accepted 15 February 2022; Published 17 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Guanyao Lu 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 recent years, with the rise of artificial intelligence, deep neural network models have been used in various image recognition researches. Land desertification is a major environmental problem facing the world at present, and how to do a good job in dynamic monitoring is particularly important. For remote sensing images, this paper constructs a GA-PSO-BP analysis model based on BP neural network, genetic algorithm, and particle swarm algorithm and compares the classification training accuracies of the four models of BP, GA-BP, PSO-BP, and GA-PSO-BP; GA-PSO-BP was selected for dynamic analysis of desertification images, and the results showed the following: (1) By comparing the regional classification training accuracies of the four models of BP, GA-BP, PSO-BP, and GA-PSO-BP, the GA-PSO-BP neural network remote sensing image classification method proposed in this paper is simple and easy to operate. Compared with traditional remote sensing image classification methods and traditional neural network classification methods, the classification accuracy of remote sensing effects is improved. (2) Carrying out desertification analysis on remote sensing images of Horqin area, from 2010 to 2015, the desertified land area in the test area increased by 1.56 km²; from 2015 to 2020, the desertified land area in the test area decreased by 1.131 km², and the desertified land in the test area from 2010 to 2020 showed a trend of increasing first and then decreasing, which is consistent with the actual situation. The GA-PSO-BP remote sensing image classification model has a good performance portability.

1. Introduction

In the early 1970s, the concept of desertification was put forward. Land desertification is also called land degradation, which is caused by the fragility of the natural environment and unreasonable human production activities. Natural factors for the rapid spread of desertified land include drought and lack of water, uncontrolled grazing destroying natural vegetation, wind erosion, and water erosion; human factors include excessive logging and waste of water resources. Once the land shows signs of desertification, the most obvious manifestation is the gradual decline or even loss of soil productivity. The phenomenon of land desertification occurs on a global scale, affecting about 25% of the world's land [1].

The prerequisite for desertification land management is real-time and accurate judgment and quantitative analysis of desertification land changes. The two most important links

are land classification and land change monitoring. Using remote sensing data and remote sensing technology to classify land is currently the main land use classification method [2]. The two methods of BP neural network and convolutional neural network also belong to the category of remote sensing supervision classification [3]. Among them, the experimental effect of BP neural network in the classification of standard remote sensing images and high score images [4], compared with traditional methods, is better in terms of accuracy. Han et al. [5] and Anwer et al. all borrowed the idea of BP neural network to classify the heterogeneous images in the study area; Han et al. built a new classification based on the classification results of the BP neural network. The adaptive classifier improves the classification accuracy; Anwer et al. [6] used the BP neural network of RBF to classify the TM images in the study area. In the research process, Park et al. [7] tried to use two different data source processing methods, perfusion and

postfusion, to explore whether the data source has an impact on CNN classification. At the same time, in terms of change monitoring, remote sensing technology has many advantages. Remote sensing images contain a large amount of information, a wider range of investigations, and a faster survey work. Through the comparative research of multi-temporal information, dynamic monitoring and trend prediction can be realized. Tasgetiren et al. [8] used differential DEM algorithm to monitor the topographic changes and topographic changes of the study area. Haque and Basak [9] combined the change vector analysis (CVA) method with the postclassification method and applied it to 4 datasets, and the accuracy was higher than the change monitoring accuracy of the algebraic method alone. Lal et al. [10] proposed a multi-time-domain remote sensing image change monitoring method based on sparse fusion and constrained k -means clustering was proposed. It adopts the idea of early fusion and later classification and compares the fusion monitoring results after classification, and the accuracy is improved.

This paper uses multispectral remote sensing images as the experimental data source, uses different classification methods, and introduces genetic algorithm (GA) and particle swarm optimization (PSO) algorithm to classify images in the study area and test area. The experimental results are compared with each other and the structure selected is simpler and the simulation error is small. The classification method is used to classify the images in the test area. Finally, the difference method is used to monitor the changes of the classified images, so as to achieve the purpose of macro-controlling the overall change trend of the desertified land [11]. The method proposed in this paper can effectively make up for the shortcomings of traditional change monitoring methods and provide a new method and idea for desertification land change monitoring.

2. Related Work

2.1. BP Neural Network. BP neural network is a self-adaptive nonlinear dynamic network system composed of neuron models, thresholds, activation functions, and so forth connected to each other [12]. The structure of a single neuron model is shown in Figure 1.

Among them, the weight of the input signal is represented by W , the bias is represented by b , and $G(\cdot)$ represents the activation function. The common feedforward neural network structure is divided into single layer and double layer, as well as recursive formula [13]. The basic structure of a multilayer feedforward neural network is from front to back: input layer (one layer) \rightarrow hidden layer (1 or more layers) \rightarrow output layer (one layer). BP neural network belongs to the category of multilayer feedforward network. The input signal conforms to the principle of forward propagation, and the error generated in the training process conforms to the principle of backpropagation. It is widely used in the field of remote sensing image processing. The training error backpropagation, in short, takes the specified training error as the target, and the training error that does not meet the target is continuously adjusted and updated

from the output layer to the input layer, until the target error is reached and the iterative update is stopped. The complete structure of the BP neural network model is shown in Figure 2. Two parts are described in detail below: signal forward propagation and error backward propagation.

Among them are the input signals (i_1 and i_2), the connection threshold between the first layer and the second layer (b_1 and b_n), the neurons in the hidden layer (h_1 and h_2), and the neurons in the output layer (o_1 and o_2). When determining the parameters of the BP neural network, you need to set the learning rate. The setting of the learning rate has only a reference range. If the value is too large, the network will oscillate, which may lead to failure to converge. If the value is too small, it will also greatly increase the time required for network convergence. At the same time, the BP neural network cannot memorize all the weight thresholds used in the learning process. Each iteration will overwrite the previous record. When there are new input samples, the weight threshold matrix of BP neural network will change accordingly, disrupting the trained sample features and classification rules, interfering with network training, and reducing the efficiency of learning.

2.2. Genetic Algorithm. Genetic algorithm is an optimization algorithm that simulates biological evolution, that is, chromosome duplication and mutation in organisms [14]. The main purpose is to select and globally optimize the objective function according to the "survival of the fittest" principle. The chromosome is the basic component, representing the vector of the objective function [15]. Compared with traditional optimization methods (enumeration, heuristic, etc.), genetic algorithm is based on biological evolution and has good convergence, less calculation time, and high robustness when calculation accuracy is required. Figure 3 shows the flow chart of the genetic algorithm.

- (1) Coding: Use a suitable coding method for vector coding, and this article uses binary coding.
- (2) Create the initial chromosome community: Use the creation function and give the community size (N) to generate the initial community.
- (3) Calculate individual fitness: Determine a reasonable fitness function, and calculate the fitness value of each chromosome in the community N according to the fitness function. The fitness value is sorted from large to small and divided into three parts.
- (4) Selection: According to the calculation result of step 3, the part of the chromosome with higher fitness will be kept after copying.
- (5) Crossover: According to the calculation result of step 3, some chromosomes with moderate fitness are retained after crossover.
- (6) Mutation: According to the calculation result of step 3, some chromosomes with lower fitness will be retained after mutation.

The above steps 1–6 need to be iterated continuously, which is in line with the principles of biological genetic

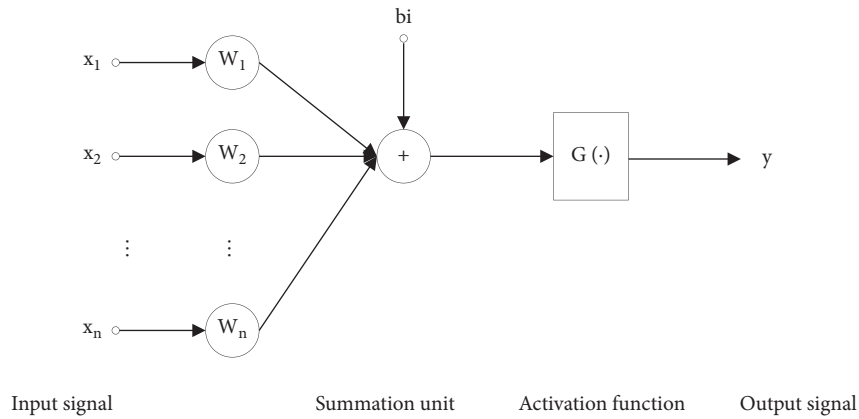


FIGURE 1: Structure diagram of a single neuron.

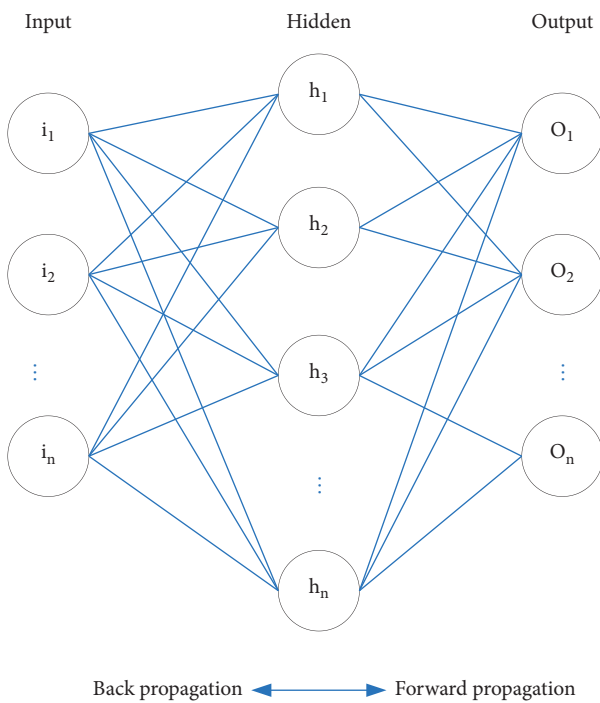


FIGURE 2: Neural network structure diagram.

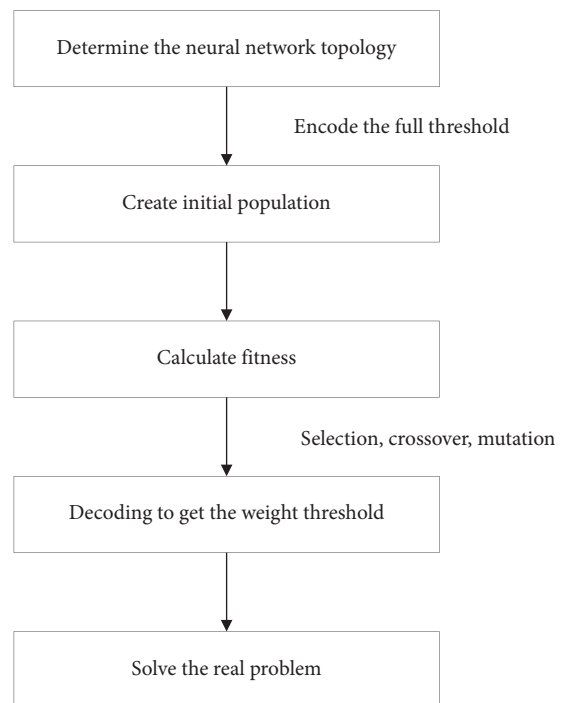


FIGURE 3: Flow chart of the genetic algorithm.

evolution. The purpose is to preserve the excellent genes in the population and inherit them to the offspring, so that the offspring have a higher average fitness and individual fitness. The genetic algorithm is simple, universal, and robust [16], which can search from different directions at the same time, which improves the global search ability to a certain extent.

2.3. Particle Swarm Algorithm. Particle swarm algorithm is also one of the commonly used global search algorithms [17]. The algorithm is proposed after studying and analyzing the predation behavior of birds [18]. The basic idea is to designate a certain point in the target area as the target food, and a flock of birds randomly search for the target food in the target area. Birds search without any prior knowledge about the target food, and it is necessary to quickly

determine the search range of the bird group closest to the target food.

Compared with other optimization algorithms, the particle swarm optimization algorithm has a simple structure and can effectively reduce the time required for network convergence [19]. At present, it is widely used in many fields such as pattern recognition, image processing, and function optimization [20]. Figure 4 shows the flow chart of particle swarm algorithm.

- (1) Randomly generate particle swarms: The size of the particle swarm and the position and velocity of the particles need to be set in advance, as well as other parameters.
- (2) Calculate the fitness value: Calculate the fitness value F of the particle based on the prescribed fitness

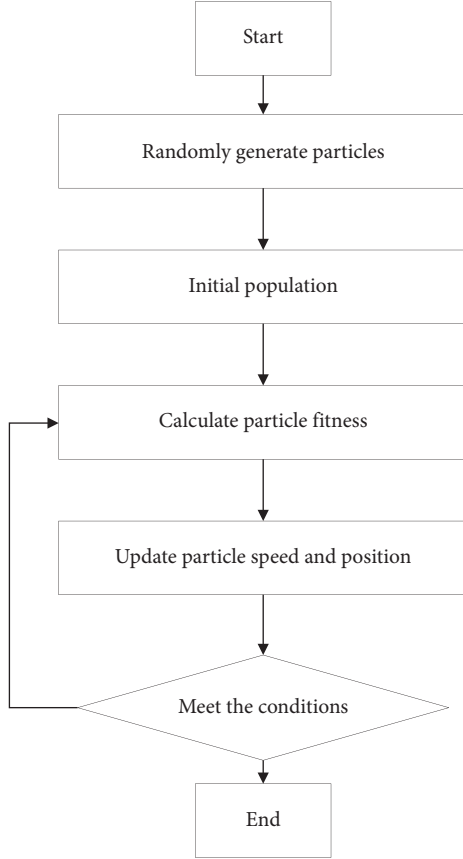


FIGURE 4: Particle swarm algorithm flow chart.

function; compare F with the individual extreme value; if $F >$ the individual extreme value, then replace the current individual extreme value with F . Globally the update of the extreme value is the same.

- (3) Update particle speed and position: Update the particles according to the two following update formulas.
- (4) Iterative judgment: If the current individual extreme value and the global extreme value meet the end conditions, it will automatically exit the iteration and complete the optimization calculation process. Otherwise, continue to iterate until the number of iterations reaches the maximum set before initialization.

The velocity update formula of a single particle in the particle swarm algorithm is as shown in formula (1):

$$\begin{aligned} \text{new_}V_i = & \omega \cdot v_i + c1 \cdot \text{rand}() \cdot (p\text{best} - P_i) \\ & + c2() \cdot \text{rand} \cdot (g\text{best} - P_i), \end{aligned} \quad (1)$$

where ω is the inertia factor, v_i is the current velocity of the particle, and P_i is the current position of the particle. The location update formula is as follows:

$$\text{new_}P_i = P_i + \text{new_}v \cdot t. \quad (2)$$

In the above equation, t is generally set to 1 by default. The PSO algorithm needs fewer parameters to adjust, and the particles have memory capabilities [21], and at the same time it can reduce the time required for network convergence.

3. GA-PSO-BP Model

This paper proposes a remote sensing image classification method and change monitoring based on a BP neural network model optimized by GA and PSO. The purpose is to effectively alleviate the instability of BP neural network memory learning, reduce the time required for network convergence, and ensure the classification accuracy of remote sensing images.

The genetic algorithm is used to further optimize the particles of the particle swarm algorithm. The optimization content includes particle speed and position, and the optimization operation includes copy, crossover, and mutation, which increases the complexity of the particle swarm algorithm and ensures that the particle swarm algorithm is executed. In the process, it will not fall into a local minimum, which ensures that the particles can complete the network optimization. Make full use of the memory characteristics of the particle swarm algorithm to ensure that the optimal solution is memorized every time in the search process. The GA-PSO-BP process is shown in Figure 5, and its specific implementation steps are as follows:

- (1) Initialization. Need to specify the number of particles of the randomly initialized population A. The remaining parameters that need to be initialized include various parameters used in the process of GA and PSO operations, as well as the maximum number of iterations N of the algorithm.
- (2) Using the particle swarm algorithm, first perform optimization on population A, and the criterion for judging whether it needs to be updated is the fitness function, and the content of the update is the particle speed and position.
- (3) The principle of particle sorting: the fitness values are arranged in order from large to small, and they are divided into three parts. Population A1 represents the part of the particles with the best fitness, population A2 represents the part of the particles with the fitness in the middle, and population A3 represents the part of the particles with the smaller fitness.
- (4) A1 (copy), A2 (crossover), and A3 (mutation).
- (5) Update gbest and ybest.
- (6) Repeat the above steps 2–5 until the number of iterations reaches the set maximum number or the objective function reaches the convergence accuracy.

Among them, in the GA-PSO iteration process, multiple parameters need to select appropriate calculation update formulas. The following is an introduction to the selection of important ones.

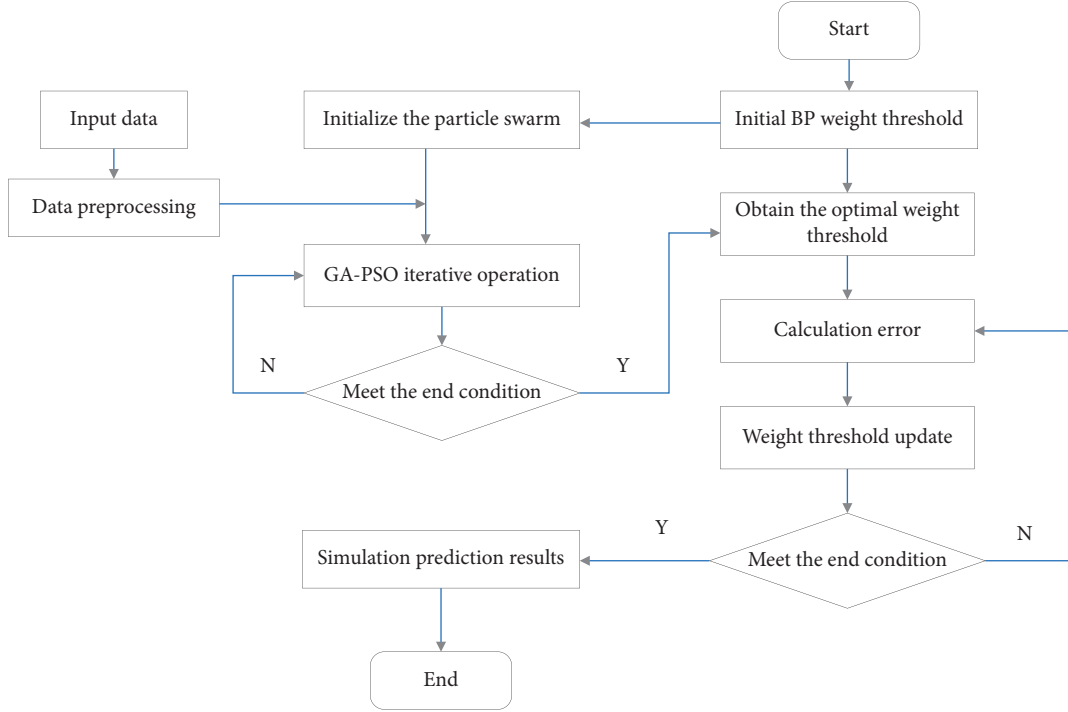


FIGURE 5: GA-PSO-BP neural network process.

For fitness function, the absolute value of the error between the predicted value of the BP neural network and the set value is used to construct the fitness function. The calculation formula is

$$f = \sum_{i=1}^n |y_i - o_i|, \quad (3)$$

where n is the total number of output nodes; y_i is the set value of the i -th node; o_i is the predicted value of the i -th node. For selection operator, the roulette method is used. For individual i in the particle swarm size N , the selection probability P_s is

$$P_s = \frac{(1/f)}{\left(\sum_{i=1}^N 1/f\right)}. \quad (4)$$

For crossover operator, equations (5) and (6) are used to perform crossover operations on the positions and velocities of particles i and j .

$$\begin{cases} X_{id}^{k+1} = \lambda_1 X_{id}^k + (1 - \lambda_1) X_{jd}^k, \\ X_{jd}^{k+1} = \lambda_1 X_{jd}^k + (1 - \lambda_1) X_{id}^k, \end{cases} \quad (5)$$

$$\begin{cases} V_{id}^{k+1} = \lambda_2 V_{id}^k + (1 - \lambda_2) V_{jd}^k, \\ V_{jd}^{k+1} = \lambda_2 V_{jd}^k + (1 - \lambda_2) V_{id}^k. \end{cases} \quad (6)$$

In formulas (5) and (6), λ_1, λ_2 is a random number in $[0, 1]$. For mutation operator, equations (7) and (8) are used to perform mutation operations on the position and velocity of particles i and j .

$$X_{id}^{k+1} = \begin{cases} X_{id}^k + (X_{id}^k - X_{\max})f(g), & \alpha_1 \geq 0.5, \\ X_{id}^k + (X_{\min} - X_{id}^k)f(g), & \alpha_1 < 0.5, \end{cases} \quad (7)$$

$$V_{id}^{k+1} = \begin{cases} V_{id}^k + (V_{id}^k - V_{\max})f(g), & \alpha_2 \geq 0.5, \\ X_{id}^k + (V_{\min} - V_{id}^k)f(g), & \alpha_2 < 0.5, \end{cases} \quad (8)$$

$$f(g) = \alpha_3 \left(1 - \frac{g}{G}\right), \quad (9)$$

where X_{\max}, X_{\min} are the upper and lower bounds of the particle position; V_{\max}, V_{\min} are the upper and lower bounds of the velocity; $\alpha_1, \alpha_2, \alpha_3$, is a random number, usually between $[0, 1]$; g represents the current evolution times; G is the largest algebra that can evolve. Inertia weight ω adopts adaptive ω adjustment method, as shown in the following equation:

$$\omega = \omega_{\max} - k \left[\frac{(\omega_{\max} - \omega_{\min})}{k} \right], \quad (10)$$

where $\omega_{\max}, \omega_{\min}$ are the upper and lower bounds of the inertia weight, k is the number of current iterations, and K is the total number of iterations.

4. Dynamic Analysis of the Impact of Desertification

4.1. Operating Environment. The experiment is based on MATLAB 2018a development platform and MATLAB 2018a neural network toolbox. All are constructed by M language. The toolbox contains a large number of functions closely

related to practical applications. When the corresponding problem needs to be solved, it can be called directly without rewriting (Table 1).

4.2. Sample Data Processing. Normalizing the input data can effectively avoid the redundancy of the data involved in the calculation. The expected output of the network corresponds to the true label of each type of feature. Before the training starts, the training samples must first be transformed into a one-dimensional vector I .

$$I_{kl} = \frac{I_{kl} - I_{\min}}{I_{\max} - I_{\min}}. \quad (11)$$

In the above equation, I_{kl} represents the pixel value of the l -th image feature in the k -th sample, and the maximum and minimum values within the range of the l -th image feature pixel value are represented by I_{\max} and I_{\min} . Based on visual interpretation and judgment of image data combined with field investigation, this paper divides the ground objects in the training area into 6 categories: water, buildings, shadows, bare land, roads, and cultivated land. Suppose that the expected output vector of the BP neural network is O :

$$O = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (12)$$

Each column of the matrix represents a category, where the expected output of the water body is $[1 \ 0 \ 0 \ 0 \ 0 \ 0]^T$, the expected output of the building is $[0 \ 1 \ 0 \ 0 \ 0 \ 0]^T$, and the expected output of the shadow is $[0 \ 0 \ 1 \ 0 \ 0 \ 0]^T$; the expected output of bare land is $[0 \ 0 \ 0 \ 1 \ 0 \ 0]^T$, the expected output of roads is $[0 \ 0 \ 0 \ 0 \ 1 \ 0]^T$, and the expected output of cultivated land is $[0 \ 0 \ 0 \ 0 \ 0 \ 1]^T$.

4.3. Model Parameter Setting

4.3.1. BP Neural Network

(1) Selection of Input and Output. The setting of the number of hidden layers and the number of nodes is the key to the network and has an impact on the learning and generalization capabilities of the network. The number of hidden layers needs to be determined through experiments or empirical rules and the principle cannot be set blindly, which may cause the network to be too computationally intensive. This may cause the network to be too computationally expensive, because each piece of sample data will have unavoidable noise, and the neural network will be affected by noise to varying degrees while learning the characteristics of the sample data. So, under normal circumstances, you can

adjust the number of hidden layer nodes instead of increasing the number of hidden layers.

When using BP neural network to classify remote sensing images, the number of bands of remote sensing data is the number of nodes in the input layer. The sample data in this article has only 3 multispectral bands; as long as it contains enough spectral feature information of the ground features, the number of nodes in the input layer is determined to be 3. The number of nodes in the input layer is determined to be 3; the number of nodes in the output layer represents the number of land types specified in the experiment, and the number of nodes in the output layer is 6, and then the number of hidden layer nodes is determined according to the prior rules of the input layer and the hidden layer given by

$$n_2 = 2 \times n_1 + 1. \quad (13)$$

It is calculated that the number of nodes in the hidden layer of the experiment in this paper n_2 is 7, and the network structure of the BP neural network is finally determined: 3-7-6 (input layer-hidden layer-output layer).

(2) Selection of Activation Function. The activation function adds nonlinear elements to the neural network. First, the activation function should satisfy the specified basic conditions. The activation function affects the time for the entire network to approach the target, and the prediction and generalization effect of the input samples. The sigmoid function is a common sigmoid function. After comparative experiments, the sigmoid function is finally determined as the network activation function in this paper.

(3) Network Global Error. Whether the global error can reach the target value is the basic criterion for judging the classification accuracy of the neural network. The setting of the global error should not only consider the actual situation of the network model but also consider the convergence speed of the network model. If the target setting is too small, the convergence speed will be reduced. It becomes very slow and cannot be used in practice; setting the target too large will lead to a substantial increase in the rate of misjudgment. Before setting the global error, specify 2-3 different global errors to conduct experiments. After the experiment, through comparison, determine the most suitable global error value. The global error TrainGoal selected in this paper is 0.000001.

(4) Learning Rate. Determining a scientific and reasonable learning rate can ensure that there is no abnormal deviation from network errors. The setting range is [0.01 0.8]. If the learning rate is set too large, the network will lose stability during the training process. If the setting is too small, it will take too long for the network to reach the target value, which is more complicated for the large number of network layers. The circumstances and the requirements for learning rate are not the same. According to the adaptive gradient descent method, the learning rate in the training process can be controlled, and it can also be automatically adjusted

TABLE 1: Introduction to commonly used functions of BP neural network.

Function category	Function name	Features
Network creation function	Newcf Newff	Create a multilayer feedforward network Create a feedforward network
Network initialization function	Initwb	Initialize the main structural parameters of the network
Transfer function	Logsig Tansig	S-type transfer function (logarithm) S-type transfer function (tangent)
Learning function	Learngd Learnqdm	Gradient descent learning function Gradient descent (momentum learning) function
Training function	Traind Adapt	Training weight/threshold Adaptive function
Simulation function	Sim Mse	Neural network simulation Mean square error function

according to the characteristics of the training situation. Finally, the adaptive gradient descent method is determined to effectively improve the performance and stability of the algorithm. Construct a BP neural network based on the above selected parameters, and perform remote sensing image classification experiments on the training area. Figure 6 shows the BP neural network remote sensing image classification process.

4.3.2. GA-BP Model. Determining the basic parameters of BP neural network is the first task for GA optimization. This experiment still uses a BP neural network whose input layer-hidden layer-output layer is “3-7-6.” In addition, determine the three most important structural parameters: training function (traingda), adaptive gradient descent method, and global error (0.000001). After determining the structural parameters of the BP neural network, the main parameters of GA need to be determined, including population size (M), maximum number of iterations (N), crossover probability (cp), and mutation probability (mp). The population size of this experiment is set to $M = 200$, the maximum number of iterations $N = 500$, the crossover probability $cp = 0.6$, and the mutation probability $mp = 0.06$. According to the parameters selected previously, the GA optimized BP neural network is constructed, and the remote sensing image classification experiment is performed on the training area. Figure 7 shows the GA optimized BP neural network remote sensing image classification process.

4.3.3. PSO-BP Model. Determining the basic parameters of BP neural network is the primary task of PSO optimization. This experiment still uses a BP neural network whose input layer-hidden layer-output layer is “3-7-6.” In addition, determine the three most important structural parameters: training function (traingda), adaptive gradient descent method, and global error (0.000001). The most important parameters for the particle swarm algorithm are the particle swarm (A), the number of particles, the learning factor ($c1$ and $c2$), the inertia weight (w), and the maximum number of iterations ($N2$). In this experiment, the number of particles in particle swarm A is set to 200, the learning factors $c1$ and $c2 = 2$, the inertia weight $w = 0.7$, and the maximum number of iterations $N2 = 500$. Construct a PSO optimized BP neural network based on the

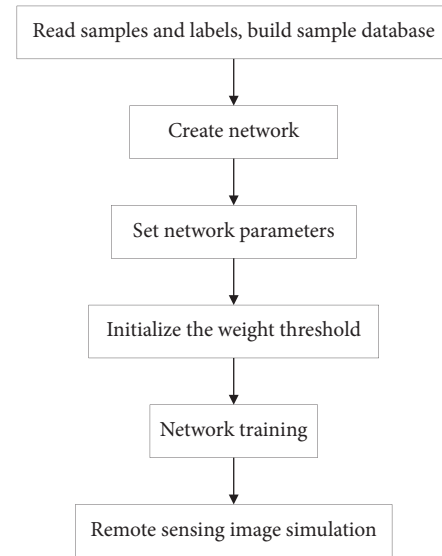


FIGURE 6: BP neural network remote sensing image classification process.

above selected parameters, and perform remote sensing image classification experiments on the training area. Figure 8 shows the PSO optimized BP neural network remote sensing image classification process.

4.3.4. GA-PSO-BP Model. Determining the basic parameters of the BP neural network is the first step in the optimization of GA combined with PSO. This experiment still uses a neural network whose input layer-hidden layer-output layer is “3-7-6.” Determine the three most important structural parameters: training function traingda, adaptive gradient descent method, and global error (0.000001).

For genetic algorithm, the population size of this experiment is $M = 200$, the maximum number of iterations is $N = 500$, the crossover probability $cp = 0.6$, and the mutation probability $mp = 0.06$. For the particle swarm algorithm, the number of particles in particle swarm A in this experiment is 200, the learning factors $c1$ and $c2 = 2$, inertia weight $w = 0.7$, and maximum number of iterations $N2 = 500$. Figure 9 shows the GA-PSO-BP neural network remote sensing image classification process, which mainly includes network creation, parameter setting, preliminary optimization of

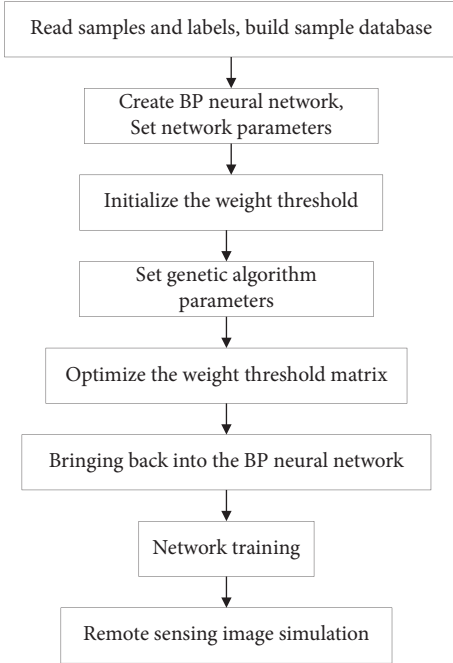


FIGURE 7: The remote sensing image classification process of the GA-BP optimization model.

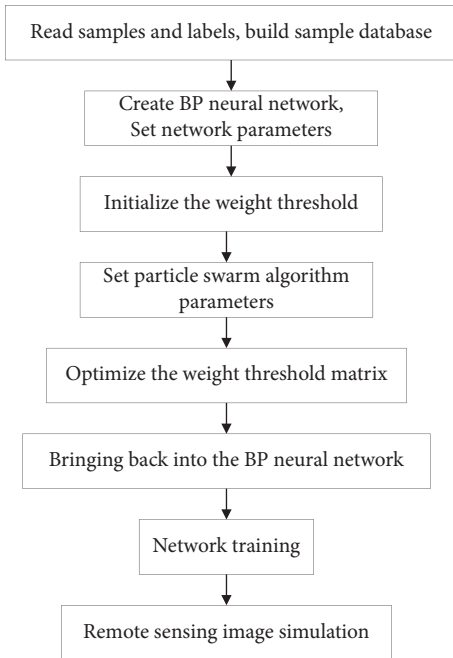


FIGURE 8: The remote sensing image classification process of the PSO-BP optimization model.

particle swarm algorithm, block and combine genetic algorithm processing according to fitness value, generation optimization, updating individual extreme value and group extreme value, network training, and remote sensing image classification.

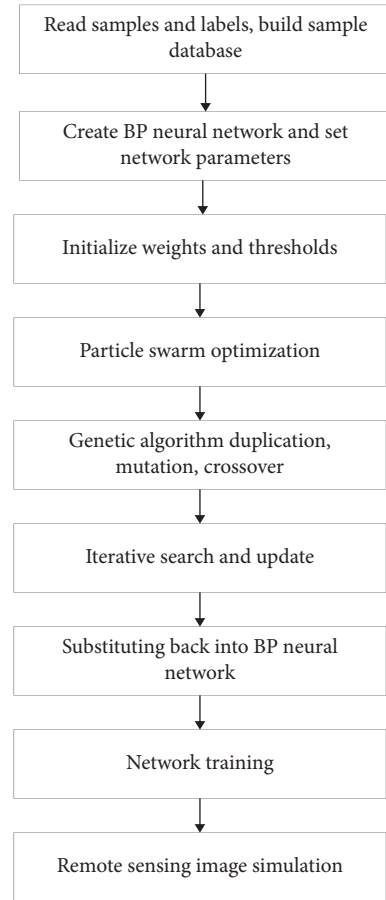


FIGURE 9: The remote sensing image classification process of the GA-PSO-BP optimization model.

4.4. Evaluation of Model Training Results. Classify and train the four models of BP neural network and GA-BP, PSO-BP, and GA-PSO-BP. This paper uses the pixel-level confusion matrix evaluation method to calculate the Kappa coefficient and the overall accuracy (OA) to classify the quality of each model. For evaluation, the real number of pixels of the same land type within the specified area and the number of pixels with a certain confidence level and the number of pixels obtained by the statistics after classification by the classification algorithm are the basic elements to form a complete confusion matrix (formula (11)). In formula (11), the columns of the matrix represent the images after classification, and the rows of the matrix represent the number of pixel categories:

$$M = \begin{bmatrix} x_{11} & \cdots & x_{m1} \\ \vdots & \ddots & \vdots \\ x_{1m} & \cdots & x_{mm} \end{bmatrix}. \quad (14)$$

In the above equation, m is the total number of categories, and x_{ij} represents the number of pixels in the study area that should belong to category i but are classified as category j . The confusion matrix calculated by Arcgis is the basis for further analysis. The main evaluation indicators of

TABLE 2: Comparison of training results of each model.

Model	Kappa	OA	Misclassification error	Omission error	Simulation error
BP	0.47	0.56	0.007	0.01	12.3158
GA-BP	0.6	0.70	0.006	0.05	11.6807
PSO-BP	0.6	0.68	0.005	0.05	11.5600
GA-PSO-BP	0.77	0.82	0.004	0.06	10.0104

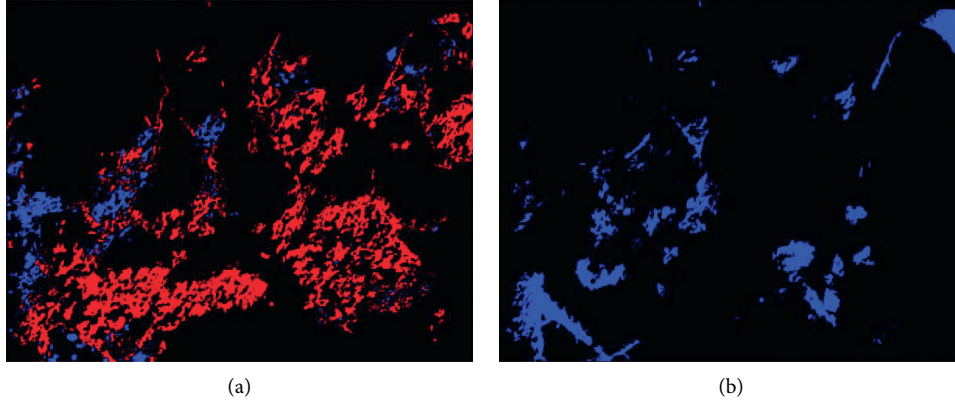


FIGURE 10: Changes in desertification land from 2010 to 2020. * Blue is decreasing area, red is increasing area, and black is nondesertification land. (a) 2010–2015. (b) 2015–2020.

this article are calculated based on the confusion matrix: OA and Kappa coefficient.

OA is calculated based on the total number of pixels that have been correctly classified, divided by the total number of pixels:

$$OA = \frac{\sum_{i=1}^m x_{ii}}{\sum_{i=1}^m \sum_{j=1}^m x_{ij}} \quad (15)$$

Kappa coefficient is calculated as

$$K = \frac{N \sum_{i=1}^m x_{ii} - \sum_{i=1}^m (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^m (x_{i+} x_{+i})} \quad (16)$$

where N represents the total number of pixels, x_{i+} represents matrix column sum, and x_{+i} represents matrix row sum.

The calculation result of the Kappa coefficient generally falls within the range of $[-1, 1]$. The specific description of the actual situation is that it needs to be distinguished by grouping the Kappa coefficients, so the Kappa coefficients are divided into 5 levels:

- (1) Falling in the range of $[0.0, 0.20]$: such a result is defined as keeping very low consistency with the actual result;
- (2) Falling within the range of $[0.21, 0.40]$: such a result is defined as maintaining general consistency with the actual result;
- (3) Falling within the range of $[0.41, 0.60]$: such a result is defined as maintaining a moderate consistency with the actual result;
- (4) Falling within the range of $[0.61, 0.80]$: such a result is defined as being highly consistent with the actual result;

- (5) Falling within the range of $[0.81, 1]$: such a result is defined as keeping almost complete with the actual result.

The classification training results of the four models are shown in Table 2.

It can be seen that the two optimized BP neural networks of GA-BP and PSO-BP have reduced simulation errors by 0.6351 and 0.7558 compared with the traditional BP neural network, respectively, and the Kappa coefficients of the two optimized BP neural networks of GA-BP and PSO-BP are the same. Both are 0.6, and the degree of agreement with the actual results is better than that of the traditional BP neural network, but the two optimized BP neural networks of GA-BP and PSO-BP are not much different from the perspective of simulation error or Kappa coefficient. It is impossible to distinguish which one is more suitable for classifying the images of the training area and the test area. For the GA-PSO-BP neural network remote sensing image classification method proposed in this article, the simulation error is significantly lower than those of the traditional BP, GA-BP, and PSO-BP neural networks. The Kappa coefficient of GA-PSO-BP neural network is 0.77.

4.5. Dynamic Analysis of Desertification Images. This article selects the Horqin area in eastern Inner Mongolia to conduct desertification image dynamic analysis. The image change monitoring method adopted is the classification post-processing method. The traditional image difference method and the implementation platform is ENVI5.3. The purpose of the difference method is to highlight the changed part of the image. To achieve the purpose of land change monitoring, Figure 10 shows the results of the change monitoring.

The red part of the figure represents the increase of desertified land from 2010 to 2020, and the blue part of the figure represents the decrease of desertified land from 2010 to 2020. Statistics show that, from 2010 to 2020, the area of desertified land in the test area increased by 296130 pixels, totaling 1.593 km², and decreased by 6,107 pixels, totaling 0.033 km². The final increase in area is 1.56 km².

Statistics show that, from 2010 to 2020, the area of desertified land in the test area decreased by 210,348 pixels, totaling 1.131 km², and there was no increase in area. From 2010 to 2020, the area of desertified land in the test area showed a trend of first increasing and then decreasing.

5. Conclusion

For remote sensing images, this paper constructs a GA-PSO-BP analysis model based on BP neural network, genetic algorithm, and particle swarm optimization algorithm and compares the classification training accuracies of the four models of BP, GA-BP, PSO-BP, and GA-PSO-BP; GA-PSO-BP was selected for dynamic analysis of desertification images, and the results showed the following:

- (1) By comparing the regional classification training accuracies of the four models of BP, GA-BP, PSO-BP, and GA-PSO-BP, the GA-PSO-BP neural network remote sensing image classification method proposed in this paper is simple and easy to operate. Compared with traditional remote sensing image classification methods and traditional neural network classification methods, the classification accuracy of remote sensing effects is improved.
- (2) Carrying out desertification analysis on remote sensing images of Horqin area, from 2010 to 2015, the desertified land area in the test area increased by 1.56 km²; from 2015 to 2020, the desertified land area in the test area decreased by 1.131 km², and the area of desertified land in the test area from 2010 to 2020 showed a trend of increasing first and then decreasing, which is consistent with the actual situation. The GA-PSO-BP remote sensing image classification model has a good performance portability.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Foundation of Guangdong Province (2015A030310505).

References

- [1] S. Kaliraj, S. M. Meenakshi, and V. K. Malar, "Application of remote sensing in detection of forest cover changes using Geo-Statistical change detection matrices – a case study of devanampatti reserve forest, tamilnadu, India," *Nature Environment and Pollution Technology*, vol. 11, no. 2, pp. 261–269, 2012.
- [2] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multi-objective genetic algorithm," *NSGA-II IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 0–197, 2002.
- [3] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy K-nearest neighbor algorithm," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 4, pp. 580–585, 1985.
- [4] S. Mirjalili, A. S. Sadiq, and H. Faris, "Genetic algorithm: theory, literature review, and application in image reconstruction," *Nature-Inspired Optimizers*, Springer, Cham, Switzerland, pp. 69–85, 2020.
- [5] M. I. N. Han, R. Quan Zhang, and M. L. Xu, "Multivariate chaotic time series prediction based on elm-plsr and hybrid variable selection algorithm," *Neural Processing Letters*, vol. 46, no. 2, pp. 705–717, 2017.
- [6] R. M. Anwer, F. S. Khan, M. Molinier, and J. Laaksonen, "Binary patterns encoded convolutional neural networks for texture recognition and remote sensing scene classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 138, pp. 74–85, 2018.
- [7] B. J. Park, H. R. Choi, and H. S. Kim, "A hybrid genetic algorithm for the job shop scheduling problems," *Computers & Industrial Engineering*, vol. 45, no. 4, pp. 597–613, 2005.
- [8] F. M. Tasgetiren, Q. K. Pan, P. N. Suganthan, and A. Oner, "A discrete artificial bee colony algorithm for the no-idle permutation flow shop scheduling problem with the total tardiness criterion," *Applied Mathematical Modelling*, vol. 37, no. 10–11, pp. 6758–6779, 2013.
- [9] M. I. Haque and R. Basak, "Land cover change detection using GIS and remote sensing techniques: a spatio-temporal study on Tanguar Haor, Sunamganj, Bangladesh," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 20, no. 2, pp. 251–263, 2017.
- [10] A. M. Lal and S. Margret Anuncia, "Semi-supervised change detection approach combining sparse fusion and constrained k means for multi-temporal remote sensing images," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 18, no. 2, pp. 279–288, 2015.
- [11] I. Kacem, S. Hammadi, and P. Borne, "Pareto-optimality approach for flexible jobshop scheduling problems: hybridization of evolutionary algorithms and fuzzy logic," *Mathematics and Computers in Simulation*, vol. 60, pp. 245–276, 2002.
- [12] M. Cerrada, G. Zurita, D. Cabrera, V. S. Rene, A. Mariano, and C. Li, "Fault diagnosis in spur gears based on genetic algorithm and random forest," *Mechanical Systems and Signal Processing*, vol. 70/71, pp. 87–103, 2016.
- [13] M. Hasan and J. M. Kim, "Fault detection of a spherical tank using a genetic algorithm-based hybrid feature pool and k-nearest neighbor algorithm," *Energies*, vol. 12, no. 6, pp. 991–1005, 2019.
- [14] A. Elazab, C. M. Wang, F. C. Jia, W. Jianhuang, L. Guanglin, and H. Qingmao, "Segmentation of brain tissues from magnetic resonance images using adaptively regularized kernel-based fuzzy c-means clustering," *Computational and mathematical methods in medicine*, vol. 2015, Article ID 485495, 12 pages, 2015.
- [15] S. Das and S. De, "Multilevel color image segmentation using modified genetic algorithm (MFGA) inspired fuzzy c-means clustering," in *Proceedings of the Second International*

- Conference on Research in Computational Intelligence & Communication Networks*, IEEE, Kolkata, India, September 2016.
- [16] M. Yazdani, M. Amiri, and M. Zandieh, "Flexible job-shop scheduling with parallel variable neighborhood search algorithm," *Expert Systems with Applications*, vol. 37, no. 1, pp. 678–687, 2010.
- [17] A. M. Baltar and D. G. Fontane, "Use of multiobjective particle swarm optimization in water resources management [J]," *Journal of Water Resources Planning and Management*, vol. 134, no. 3, pp. 257–265, 2008.
- [18] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the Icn95-international Conference on Neural Networks*, IEEE, Perth, WA, Australia, November 1995.
- [19] J. Park, G. Lee, and Y. Choi, "Determining ARMA model parameters for biomedical signal applications based on estimation of the equivalent AR model," in *Proceedings of the International Conference on Ubiquitous Robots & Ambient Intelligence*, pp. 456–457, IEEE, Goyangi, South Korea, October 2015.
- [20] Ç. Gürcan and K. Ali, "Optimization of thermodynamic performance with simulated annealing algorithm: a geothermal power plant," *Renewable Energy*, vol. 172, 2021.
- [21] C. Praveen and R. Duvigneau, "Low cos PSO using meta-models and inexact pre-evaluation: application to aerodynamic shape design," *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 9–12, pp. 1087–1096, 2009.

Research Article

Human Resource Planning and Configuration Based on Machine Learning

Shuai Yuan,¹ Qian Qi ,² Enliang Dai,³ and Yongfeng Liang⁴

¹School of Management Engineering and Business, Hebei University of Engineering, Handan 056000, China

²School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100000, China

³School of Labor and Human Resources, Renmin University of China, Beijing 100000, China

⁴Beijing Miss Fresh E-Commerce Co. Ltd, Beijing 100000, China

Correspondence should be addressed to Qian Qi; qiqian145045@163.com

Received 30 December 2021; Accepted 20 January 2022; Published 15 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Shuai Yuan 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.

Human resources are the core resources of an enterprise, and the demand forecasting plays a vital role in the allocation and optimization of human resources. Starting from the basic concepts of human resource forecasting, this paper employs the backpropagation neural network (BPNN) and radial basis function neural network (RBFNN) to analyze human resource needs and determine the key elements of the company's human resource allocation through predictive models. With historical data as reference, the forecast value of current human resource demand is obtained through the two types of neural networks. Based on the prediction results, the company managers can carry out targeted human resource planning and allocation to improve the efficiency of enterprise operations. In the experiment, the actual human resource data of a certain company are used as the experimental basic samples to train and test the two types of machine learning tools. The experimental results show that the method proposed in this paper can effectively predict the number of personnel required and can support the planning and allocation of human resources.

1. Introduction

Human resources are valuable corporate resources and are of great significance to their predictive analysis. Many experts and scholars all over the world have conducted relevant research studies [1–4]. Human resource demand forecasting generally needs to follow the principle of correlation and the principle of inertia. The principle of correlation is based on the correlation between the research objects and uses other objects to predict the targeted object [5–12]. For example, there is a clear correlation between A, B, and C. In the research process, rich data of A, B, and C can be obtained. At the same time, the trend value of B and C can be predicted by suitable forecasting methods, and finally the correctness can be achieved by making predictions based on A. The principle of inertia specifically refers to the slow progress of A or its regular development, and some valid past data can be obtained. Under this premise, you can choose appropriate means to predict the

trend value of A. The human resource demand forecasting model is mainly based on qualitative and quantitative analysis. Qualitative analysis methods mainly include the Delphi method, subjective judgment method, microintegration method, and so on. The above methods rely more on experts or experienced people, and they all have the disadvantage of subjective components. As a result, the prediction results are prone to be non-consensual judgments. The quantitative analysis methods mainly include the production function method, ratio method, trend method, regression analysis method, and so on. The above quantitative analysis is based on the existing data for predictive analysis. The prediction logic is rigorous and overcomes the shortcomings of qualitative methods. However, there are also problems such as low prediction accuracy and difficulty in collecting data. Therefore, the combination of quantitative and qualitative methods has naturally become the research direction of enterprise human resource forecasting [13–17].

Zhandong and Chi [18] used the Delphi method to study the human resource forecasting of enterprises and explained the importance of human resource forecasting in enterprise management. Belhaj et al. [1] used the Markov model to predict the human resources of the enterprise and obtained the demand for the human resources of the enterprise in the future [19]. Wu and Nagahashi [20] used the grey forecasting model to carry out human resource forecasting analysis for enterprises, which provided a reference for its human resource planning. Qu et al. [21] predicted the human resources of enterprises based on backpropagation neural network (BPNN) and believed that network can obtain better prediction results. There are many factors affecting human resource demand and non-linear correlation. At the same time, traditional analysis and forecasting methods such as the analytic hierarchy process, multiple linear regression method, and Delphi method have shortcomings and shortcomings such as low prediction accuracy and biased subjective analysis. Therefore, common predictive analysis methods cannot truly reflect the nature of the problem. As a remedy, researchers began to study prediction methods based on machine learning models, which have higher superiority compared with traditional prediction methods [21–25]. Common modern machine learning methods include BPNN, K-nearest neighbors (KNN), support vector machine (SVM), and so on. In addition, the machine learning methods based on tree models were also widely used for data prediction, such as basic decision tree models and related integrated models such as random forests. Therefore, the research on human resource forecasting using machine learning plays an important role in improving forecasting accuracy [26–30].

This paper studies the method of personnel resource planning and allocation based on machine learning. The basic idea of this method is to use machine learning algorithms to predict enterprise human resource needs and dynamically adjust staffing accordingly. Specifically, this paper uses two machine learning models, BPNN and radial basis function neural network (RBFNN). Both types of models are based on neural networks, and robust model parameters are obtained by optimizing training algorithms. For the company's human resource needs, the historical data are used as training samples to train the two models to obtain a robust prediction model. Finally, the corresponding predicted value of human resource demand can be obtained under current conditions, which can be used as a reference for enterprise management personnel in personnel planning and configuration. The experiment uses human resource data of a certain company as the samples. Also, the validity of the proposed method can be verified according to the experimental results.

2. Machine Learning Models

2.1. BPNN. BPNN is a kind of multilayer feedforward neural network with signal forward transmission and error backward propagation [11–15]. It is widely used in the fields of function approximation and pattern recognition. The

topological structure of BPNN includes input layer, hidden layer, and output layer, and each layer is composed of neuron connections. In the forward calculation process of the BPNN, information is input by the input layer and processed and calculated by the hidden layers, and the output layer outputs the processing results. In the backward propagation of the error, the error of the processing result of the output layer is calculated. The error signal is propagated back, and the weight of the connection between each neuron is corrected by the method of gradient descent, so as to realize the network optimization.

The training of BPNN includes two processes: forward propagation of signal and backward feedback of error. The forward propagation means that in the calculation, the signal enters the hidden layer after parallel weighting calculation from the input layer and then enters the output layer through weighting processing to obtain the output. The direction of the reverse feedback process is carried out from output to input. The weights and thresholds are adjusted according to the actual error. Through continuous repetition of forward calculation and reverse feedback, the output result can finally meet the requirements. A basic model of BPNN is shown in Figure 1. In this figure, $(x_1, x_2, \dots, x_j, \dots, x_m)$ is the input quantity, and the subscript is the input quantity number, which corresponds to the input layer node; $(y_1, y_2, \dots, y_k, \dots, y_t)$ is the output quantity, and the subscript is the output quantity number, which corresponds to the output layer node; $(\theta_1, \theta_2, \dots, \theta_i, \dots, \theta_q)$ are the thresholds introduced for the hidden layer, and the subscript is the hidden node number; there may be multiple hidden layers in the neural network; $(a_1, a_2, \dots, a_k, \dots, a_t)$ are the threshold values introduced for the output layer, and the subscript is the output node number. There is only one output layer in the neural network; ε denotes the error value.

In the process of forward propagation, the input n_i and output o_i of the i th node in the hidden layer can be calculated as follows:

$$n_i = \sum_{j=1}^m w_{ij}x_j + \theta_i, \quad (1)$$

$$o_i = \phi(\text{net}_i) = \phi(n_i).$$

In the process of forward propagation, the input net_k and output of the k th in the output layer are calculated as follows:

$$\text{net}_k = \sum_{u=1}^q w_{ku}o_u + a_k = \sum_{u=1}^q w_{ku}\phi\left(\sum_{j=1}^m w_{uj}x_j + \theta_u\right) + a_k, \quad (2)$$

$$\text{out}_k = \psi(\text{net}_k),$$

where w_{ij} is the weight of the j th input variable at the i th node of the hidden layer; ϕ is the activation function of the hidden layer; and ψ is the activation function of the output layer. In order to enhance the large-scale non-linear fitting ability, according to actual use experience, the hidden layer activation function selects the bipolar S-shaped tansig function, and the output layer activation function selects the linear purelin function.

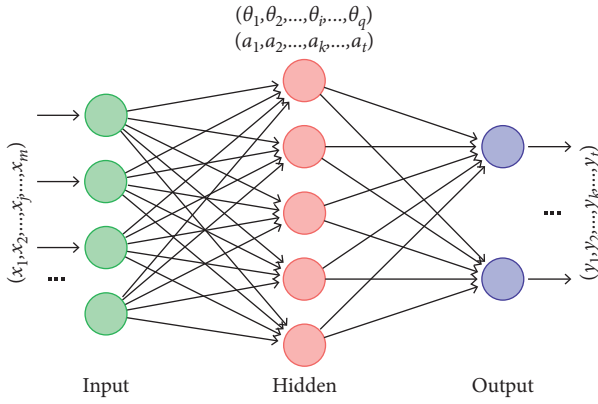


FIGURE 1: Basic structure of BPNN.

The error function E_p of the P th sample is shown in equation (3), where T_k is the expected output value of the k th node.

$$E_p = \frac{1}{2} \sum_{k=1}^l (T_k - \text{out}_k)^2. \quad (3)$$

The total error E of P training samples is

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^l (T_k^p - \text{out}_k^p)^2. \quad (4)$$

In the feedback process, the weight correction amount and threshold correction amount of the hidden layer and the output layer can be written as

$$\begin{aligned} \Delta w_{ij} &= -\eta \frac{\partial E}{\partial w_{ij}}, \Delta \theta_i = -\eta \frac{\partial E}{\partial \theta_i}, \\ \Delta w_{kj} &= -\eta \frac{\partial E}{\partial w_{kj}}, \Delta a_k = -\eta \frac{\partial E}{\partial a_k}, \end{aligned} \quad (5)$$

where η is the coefficient determining adjustment rate.

The specific steps of BPNN are as follows:

Step 1: Network initialization: the necessary network parameters are determined. Generally, there are the number of input layer nodes, the number of hidden layer nodes, the number of output layer nodes, connection weights and thresholds, transfer function types, and so on.

Step 2: According to the parameters determined in Step 1, the hidden layer output calculation is carried out.

Step 3: Same as above, the input calculation of the output layer is carried out.

Step 4: The network error is calculated as follows: the expected output-network predicted output.

Step 5: The connection weight and threshold are updated according to the network error and network learning rate.

Step 6: The judgment is made on whether the algorithm is terminated. If it is not over, return to Step 2 to continue network training.

It should be noted that there are many kinds of transfer functions, and the threshold transfer function (Hardlim) is generally used. The second is that the network data need normalized data during the training process, which requires that the data are normalized to (0, 1). Also, the data are restored when the data are output. Third, in general, the number of hidden layer nodes needs to be determined manually. Set the number of input layer nodes to be n , the number of hidden layer nodes to be l , and the number of output layer nodes to be m , and they satisfy the following relationship:

$$\begin{cases} l < n - 1, \\ l < \sqrt{m+n} + a, \quad (0 < a < 10), \\ l = \log_2 n. \end{cases} \quad (6)$$

2.2. RBFNN. Furthermore, because the BPNN generally needs to iteratively determine the connection weights, a relatively large time delay will be generated for a large amount of data processing [16–20]. The RBFNN obtains a value by calculating the norm of the input sample and the hidden layer point (center point) and substituting it into the radial basis function (Gaussian function, quadratic function, inverse quadratic function, and so on). After the weights are multiplied and added, the corresponding output is obtained. The network is simple, and the learning convergence speed is faster, which can make up for the deficiencies of BPNN.

RBFNN is a three-layer forward network, and its network structure is shown in Figure 2. Among them, W_1 and b_1 are the connection weight matrix and bias vector from the input layer to the hidden layer. W_2 and b_2 are the connection weight matrix and the connection weight matrix from the hidden layer to the output layer. The input layer is composed of signal nodes, and the number of neurons is the dimension of the input sample. The activation function of the hidden layer neurons is a RBF that is radially symmetric and attenuated to the center point and is commonly used as Gauss function, reflected sigmoidal function, inverse multiquadric function, etc. The output layer responds to the input pattern, and the number of neurons is equal to the dimension of the output sample. The transformation from the input space to the hidden layer space is non-linear, and the transformation from the hidden layer space to the output layer space is linear, so the mapping of RBFNN from input to output is also non-linear.

The basic idea of RBFNN is to use RBF as the activation function of the hidden layer neurons to form the hidden layer space, so that the input vector can be directly mapped to the hidden layer space. Also, the mapping from the hidden layer space to the output layer space is linear, that is, the network output of RBFNN is the linear weighted sum of the output of hidden layer neurons. The connection weight from the hidden layer to the output layer is an adjustable parameter of RBFNN. It can be seen that although the network input to output mapping is non-linear, the network output is linear for the adjustable parameters, so the adjustable parameters of the network can be solved directly by linear

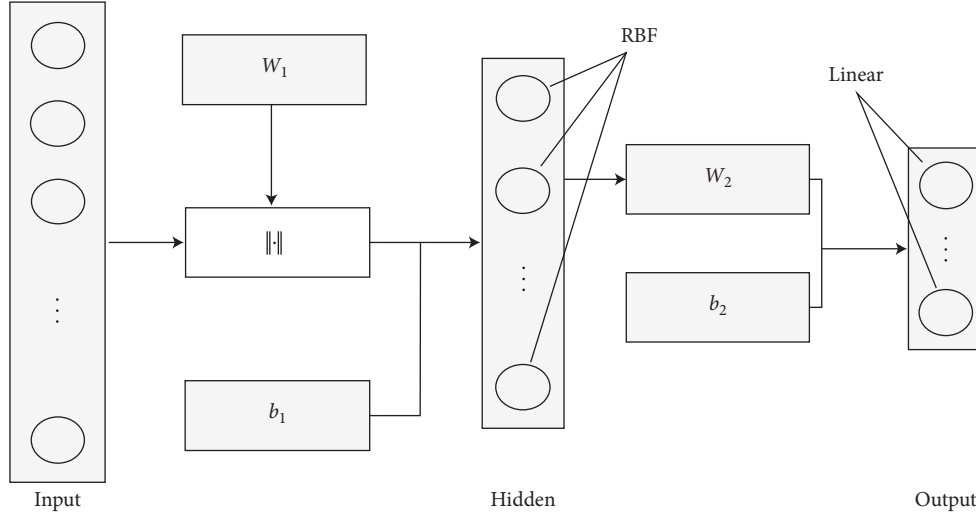


FIGURE 2: Basic structure of RBFNN.

equations, which greatly accelerates the learning speed and avoids the local minima problems. RBFNN is a function approximation network that non-linearly maps the input space to the output space. The weight vector of the array beam design is a non-linear function of the position of the array element, so the use of RBFNN can realize the mapping from the position of the array element to the weight vector of the array.

According to the above discussion, the overall expression of RBFNN is

$$f(x) = \omega_0 + \sum_{i=1}^n \omega_i \varphi(\|x - c_i\|), \quad (7)$$

where ω is the connection weight, which can be iteratively obtained by the least squares method, and the calculation formula is

$$\omega = \exp\left(\frac{h}{c_{\max}^2} \|x_p - c_i\|^2\right), \quad i = 1, 2, \dots, h, p = 1, 2, \dots, P. \quad (8)$$

The specific learning algorithm steps are as follows:

Step 1: The K-means clustering method is used to solve the center of the radial basis function; generally, after the network initialization, the input data are calculated according to the nearest neighbor rule, and grouping and re-adjusting the clustering center are completed in 3 steps.

Step 2: The variance of RBF of the RBFNN is calculated.

Step 3: The least squares algorithm is employed to calculate the weight between the hidden layer and the output layer.

3. Experiment and Analysis

3.1. Dataset and Comparison Method. Taking a company's personnel data as an example, the original data (2009–2020) are divided into two parts. The first 10-year data are the basic

data, and the 11th year data are the verification data, and the model is trained and solved. When judging the pros and cons of different models' predictive performance, some quantitative index systems are needed. According to the existing literature, this paper selects three indicators: mean squared error (MSE), mean absolute percentage error (MAPE), and symmetric mean absolute percentage error (SMAPE), as the evaluation indexes for the predictive performance. The specific calculation formulas of the above three error indicators are as follows:

$$\begin{aligned} \text{MSE} &= \frac{1}{N} \sum_{\tau=t_0}^{t_{N-1}} (y_{\tau} - y_{\tau}^0)^2, \\ \text{MAPE} &= \frac{100\%}{N} \sum_{\tau=t_0}^{t_{N-1}} \left| \frac{y_{\tau} - y_{\tau}^0}{y_{\tau}^0} \right|, \\ \text{SMAPE} &= \frac{100\%}{N} \sum_{\tau=t_0}^{t_{N-1}} \frac{|y_{\tau} - y_{\tau}^0|}{(|y_{\tau}| + |y_{\tau}^0|)/2}. \end{aligned} \quad (9)$$

4. Results and Analysis

In addition to the two models used in this paper, two traditional methods from the literature [12, 18] are also selected for comparative analysis in the experiment. Taking selected experimental data as samples, based on the three indicators of MSE, MAPE, and SMAPE to test different methods, the statistics of the results of different methods are shown in Table 1. It can be seen that the overall performance of the BPNN and RBFNN selected in this paper is better than that of the two traditional methods, reflecting the performance advantages of the machine learning models. Comparing BPNN and RBFNN, the latter has more advantages in overall performance because of the consideration of non-linear factors.

In the actual process, due to the influence of external factors such as market changes and changes in the

TABLE 1: Comparison of different methods in the prediction of human resources.

Method	Evaluation index		
	MSE	MAPE	SMAPE
BPNN	0.32	0.43%	0.43%
RBFNN	0.41	0.47%	0.46%
Comparison 1	0.45	0.52%	0.57%
Comparison 2	0.47	0.54%	0.52%

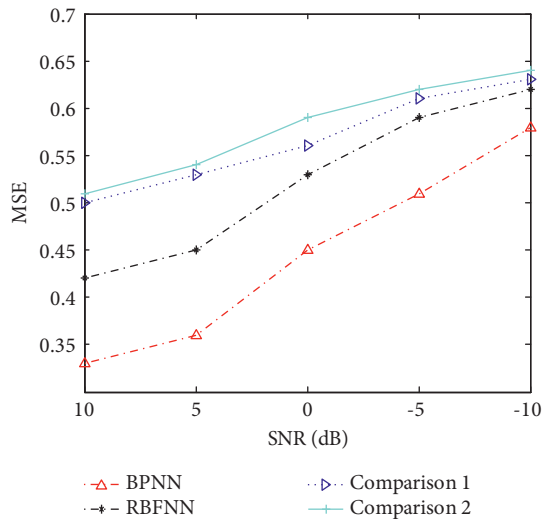


FIGURE 3: Performance measured by MSE achieved by different methods under noises.

international situation, the prediction model may have certain deviations. For this reason, this paper applies a certain degree of noise conditions to the experimental data to reflect the influence of external influences on the allocation of human resources. On this basis, the performance trend of various methods is tested using MSE as the basic evaluation index, and the results are shown in Figure 3. It can be seen from the figure that the performance of various methods is degraded to a certain extent due to the influence of noise. As the signal-to-noise ratio (SNR) decreases, the MSE keeps increasing. In comparison, the two models used in this paper can maintain the best prediction performance under different noise interference conditions, which further enhances its performance advantages. Comparing BPNN and RBFNN, the latter has better overall noise robustness due to the consideration of the possible non-linear effects of noise.

5. Conclusion

The sustainable demand forecast of human resources is the prerequisite and basis for the correct deployment of personnel planning. This paper proposes a human resource prediction method based on machine learning to address the above problems. Two types of neural networks, BPNN and RBFNN, are used to predict the human resource needs of enterprises. We train the two types of models based on the historical data of human resources to obtain predictive

models. According to the current enterprise situation, the current enterprise human resource forecast value can be obtained. In the experiment, the two models are tested and verified using human resource data of a certain enterprise. The experimental results show that the proposed method is effective for enterprise personnel resource forecasting and can support enterprise managers to carry out scientific personnel planning and allocation.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] R. Belhaj and M. Tkiouat, "A Markov model for human resources supply forecast dividing the HR system into subgroups[J]," *Journal of Service Science and Management*, vol. 6, no. 3, pp. 211–217, 2013.
- [2] Y. Wu, Z. Wang, and S. Wang, "Human resource allocation based on fuzzy data mining algorithm," *Complexity*, vol. 2021, p. 11, Article ID 9489114, 2021.
- [3] H. Y. Chiang and B. M. T. Lin, "A decision model for human resource allocation in project management of software development," *IEEE Access*, vol. 8, pp. 38073–38081, 2020.
- [4] Q. Feng, Z. Feng, and X. Su, "Design and simulation of human resource allocation model based on double-cycle neural network," *Computational Intelligence and Neuroscience*, vol. 2021, p. 10, Article ID 7149631, 2021.
- [5] B. Memishi, M. S. Pérez, and G. Antoniu, "Feedback-based resource allocation in MapReduce-based systems," *Scientific Programming*, vol. 2016, p. 13, Article ID 7241928, 2016.
- [6] N. Alhussien and T. A. Gulliver, "Optimal resource allocation in cellular networks with H2H/M2M coexistence," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 12951–12962, 2020.
- [7] G. Kakkavas, K. Tsitseklis, V. Karyotis, and S. Papavassiliou, "A software defined radio cross-layer resource allocation approach for cognitive radio networks: from theory to practice," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, pp. 740–755, 2020.
- [8] F.-H. Tseng, X. Wang, L.-D. Chou, H.-C. Chao, and V. C. M. Leung, "Dynamic resource prediction and allocation for cloud data center using the multiobjective genetic algorithm," *IEEE Systems Journal*, vol. 12, pp. 1688–1699, 2018.
- [9] S. Basu, S. Roy, S. Bandyopadhyay, and S. Das Bit, "A utility driven post disaster emergency resource allocation system using DTN," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 7, pp. 2338–2350, 2020.
- [10] A. Aijaz, "Toward human-in-the-loop mobile networks: a radio resource allocation perspective on haptic communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4493–4508, 2018.
- [11] A. Masmoudi, K. Mnif, and F. Zarai, "A survey on radio resource allocation for V2X communication," *Wireless Communications and Mobile Computing*, vol. 2019, p. 12, Article ID 2430656, 2019.

- [12] N. Arora and P. D. Kaur, "A Bolasso based consistent feature selection enabled random forest classification algorithm: an application to credit risk assessment," *Applied Soft Computing Journal*, vol. 2020, pp. 1–15, 2020.
- [13] D. P. Fonseca, P. F. Wankea, and H. L. Correa, "A two-stage fuzzy neural approach for credit risk assessment in a Brazilian credit card company[J]," *Applied Soft Computing Journal*, vol. 2020, pp. 1–13, 2020.
- [14] J. Jin Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 3, pp. 299–310, 2005.
- [15] D. D. Wu, D. L. Olson, and C. Luo, "A decision support approach for accounts receivable risk management," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 12, pp. 1624–1632, 2014.
- [16] N. Wan, L. Li, C. Ye, and B. Wang, "Risk assessment in intelligent manufacturing process: a case study of an optical cable automatic arranging robot," *IEEE Access*, vol. 7, pp. 105892–105901, 2019.
- [17] Z. Qu, "Application of improved PCA in risks assessment technology of enterprise information security," in *Proceedings of the 2009 2nd International Conference on Power Electronics and Intelligent Transportation System (PEITS)*, pp. 58–61, Shenzhen, China, December 2009.
- [18] X. Zhandong and G. Chi, "Bank-enterprise project risk assessment model based on the information entropy method," in *proceedings of the 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, pp. 998–1001, Deng Feng, China, August 2011.
- [19] P.-B. Zhang and Z.-X. Yang, "A novel AdaBoost framework with robust threshold and structural optimization," *IEEE Transactions on Cybernetics*, vol. 48, no. 1, pp. 64–76, 2018.
- [20] S. Wu and H. Nagahashi, "Parameterized AdaBoost: introducing a parameter to speed up the training of real AdaBoost," *IEEE Signal Processing Letters*, vol. 21, no. 6, pp. 687–691, 2014.
- [21] Q. Qu, C. Liu, and X. Bao, "E-commerce enterprise supply chain financing risk assessment based on linked data mining and edge computing," *Mobile Information Systems*, vol. 2021, p. 9, Article ID 9938325, 2021.
- [22] T. Chen, Q. Yang, Y. Wang, and S. Wang, "Double-layer network model of bank-enterprise counterparty credit risk contagion," *Complexity*, vol. 2020, p. 25, Article ID 3690848, 2020.
- [23] Y. Shao, X. Yao, L. Tian, and H. Chen, "A multiswarm optimizer for distributed decision making in virtual enterprise risk management," *Discrete Dynamics in Nature and Society*, vol. 2012, p. 24, Article ID 904815, 2012.
- [24] S. Dong and C. Liu, "Sentiment classification for financial texts based on deep learning," *Computational Intelligence and Neuroscience*, vol. 2021, p. 9, Article ID 9524705, 2021.
- [25] X. Wei, "A method of enterprise financial risk analysis and early warning based on decision tree model," *Security and Communication Networks*, vol. 2021, p. 9, Article ID 6950711, 2021.
- [26] J. Wu, C. Li, and Y. Huo, "Safety assessment of dangerous goods transport enterprise based on the relative entropy aggregation in group decision making model," *Computational Intelligence and Neuroscience*, vol. 2014, p. 7, Article ID 571058, 2014.
- [27] A. Galassi, J. D. Martín-Guerrero, E. Villamor, C. Monserrat, and M. J. Rupérez, "Risk assessment of hip fracture based on machine learning," *Applied Bionics and Biomechanics*, vol. 2020, p. 13, Article ID 8880786, 2020.
- [28] B. Williams, B. Allen, Z. Hu et al., "Real-time fall risk assessment using functional reach test," *International Journal of Telemedicine and Applications*, vol. 2017, p. 8, Article ID 2042974, 2017.
- [29] Y. Jianxing, C. Haicheng, W. Shibo, and F. Haizhao, "A novel risk matrix approach based on cloud model for risk assessment under uncertainty," *IEEE Access*, vol. 9, pp. 27884–27896, 2021.
- [30] B. R. Greene, S. J. Redmond, and B. Caulfield, "Fall risk assessment through automatic combination of clinical fall risk factors and body-worn sensor data," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 725–731, 2017.

Research Article

English Text Readability Measurement Based on Convolutional Neural Network: A Hybrid Network Model

Lihua Jian ¹, Huiqun Xiang,^{2,3} and Guobin Le^{2,3}

¹School of International Education, Hunan University of Medicine, Hunan, Huaihua 418000, China

²Changsha Vocational and Technical College, Hunan, Changsha 410200, China

³School of Foreign Languages, Huaihua University, Huaihua 418000, China

Correspondence should be addressed to Lihua Jian; 202110080102@hunnu.edu.cn

Received 13 January 2022; Revised 25 January 2022; Accepted 9 February 2022; Published 15 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Lihua Jian 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.

Text readability is very important in meeting people's information needs. With the explosive growth of modern information, the measurement demand of text readability is increasing. In view of the text structure of words, sentences, and texts, a hybrid network model based on convolutional neural network is proposed to measure the readability of English texts. The traditional method of English text readability measurement relies too much on the experience of artificial experts to extract features, which limits its practicability. With the increasing variety and quantity of text readability measurement features to be extracted, it is more and more difficult to extract deep features manually, and it is easy to introduce irrelevant features or redundant features, resulting in the decline of model performance. This paper introduces the concept of hybrid network model in deep learning; constructs a hybrid network model suitable for English text readability measurement by combining convolutional neural network, bidirectional long short-term memory network, and attention mechanism network; and replaces manual automatic feature extraction by machine learning, which greatly improves the measurement efficiency and performance of text readability.

1. Introduction

As long as people create, study, share, and disseminate ideas through written language, the concept of text difficulty will be always an important aspect of people's communication and education [1–3]. More than 2000 years ago, ancient Athenian scholars paid attention to the difficulty of reading the text and pointed out that students studying law usually face the problem that their laws are difficult to be understood by the audience. If the audience does not understand them, they cannot understand and support the truth they say in the legal debate. In order to better explain the meaning of the law to the audience, language rhetoric training has become an important part of learning law. In the process of language learning, improving reading ability is also an important part of language acquisition, and its reading comprehension ability is also the key standard to evaluate learners' language skills [2]. The most effective way to improve reading comprehension is to provide learners with reading materials slightly higher than their reading ability.

Reading too simple text is meaningless repetitive work. If the text is too difficult, language learners will lose their confidence and interest in learning the language.

Therefore, classifying reading materials through the reading ability of learners or audiences plays a vital role in foreign language learning [3, 4]. This task of sorting out reading learning materials according to the reading difficulty of the text is called text readability measurement or text reading difficulty evaluation, which is an important natural language processing (NLP) problem [5–8]. Since the last century, there has been a systematic and scientific method for understanding the subjective and objective factors related to text readability, better supporting readers to understand more difficult texts, or correctly finding the task of text reading difficulty. Based on the research of these systems, text readability has been defined as the sum of all elements that affect readers' understanding of text materials, reading speed, and interest in text content. These elements may include the complexity of text syntax, the semantic

familiarity of readers with some concepts in the text, whether there are supporting graphics or illustrations to explain the text, the complexity of logical argument or inference used to connect various views in the text, and many other important contents [9–11]. In addition to these text features, readers' own characteristics, that is, their education, social background, interests, professional knowledge, learning motivation, and other factors, can play a key role in the readability of the text.

In view of the importance of text readability in meeting people's information needs and the explosive growth of modern information, the measurement demand of text readability is increasing, and the influence of effective text readability evaluation is also increasing day by day [12]. The so-called effectively quantifying the reading difficulty of a text means that the reading difficulty level of the corresponding text is expressed by using the text as input and in the form of prediction such as estimated digital score or difficulty level category label, or used to indicate the reading level and understanding ability of a given population to the corresponding text [13–17]. In this paper, we focus on the internal language feature factors affecting text readability, such as semantics and syntax, but do not pay attention to the external feature factors affecting text reading difficulty, such as font size or font color contrast, which affect readers' visual decoding ability, as well as readers' educational background, interests, and hobbies and other factors related to readers themselves.

The significance of this paper lies in the following points.

- (1) Automatic and effective measurement of text readability can liberate the unnecessary labor of some people, such as teachers, students, and web text processors, in finding and classifying the reading difficulty of relevant texts
- (2) Automatic and effective measurement of text readability is of great significance for the accessibility of key information and also plays a key role in specific application fields

The functions of readability measurement include the following: (1) It can provide language learners with extra-curricular reading materials of different levels of difficulty suitable for their reading abilities at different stages, such as graded reading. (2) It can provide language teachers with teaching resources suitable for the difficulty of reading and provide guidance for their application of compiling teaching materials and test questions. (3) It can automatically calibrate and simplify public and private health information so that the public and patients can read and understand medical related text resources such as medical instructions, health questionnaires, and brochures. (4) It can provide suggestions for businesses to make effective product guides and other text documents for the public. (5) It can also be further applied to the accurate retrieval and recommendation of web text.

2. Related Works

The measurement of text readability usually refers to determining the difficulty of text content being understood by people [16, 17]. Generally, the readability of a text can be

measured by a predefined readability level or readability score. In this paper, the readability level is used to measure the readability of text. The measurement of text readability can be regarded as a classification problem, that is, how to learn the prediction model according to the text set with determined readability level and then use the model to predict the text with unknown readability level.

The research on text readability measurement has a history of at least one century. However, this is far from a "solved" problem, and the automatic measurement of text readability is still a challenging research field. The research on the measurement of readability can be traced back to the 1920s. Early readability studies mainly focused on the lexical factors of the text and used proxy variables to represent the relevant characteristics of vocabulary, such as the difficulty, diversity, and scope of use. Whether one vocabulary difficulty standard is better than the other depends mainly on the experience of expert judges and correlation analysis. These works showed that the research on text readability began to pay attention to all aspects of feature selection. From 1940s to 1990s, the readability research system was initially formed. During this period, researchers continued to try various readability formulas, introduce proxy variables of lexical and syntactic information into the formulas, and make linear combinations, hoping to accurately evaluate the text readability and obtain an optimal reading difficulty measurement standard [18–20].

From 1980s to 1990s, researchers began to pay attention to the structural information of text and introduced cognitive theories such as connection theory, conceptual schema theory, prototype theory, and diffusion activation theory into the field of text readability to explain the way people store and retrieve information in long-term memory. Through human cognitive style, the concept of text readability is associated with text structure, and the characteristics of text organization structure, discourse coherence, and cohesion are introduced [21]. At the same time, we also pay attention to the measurement of lexical features, introduce statistical language model to statistically analyze the words and word occurrence order in a given text set, count the occurrence frequency of different words or word combinations in the text set, and use this probability to measure the difficulty of vocabulary reading.

Statistical language model is applied to measure the readability of science and technology web pages [22]. After that, with the development of natural language processing technology, such as part-of-speech tagging, syntax analysis, and language model, researchers can more deeply mine the content and structure of text, which makes the research of readability have new progress. Then, new text features are constantly explored, and new theories such as information theory have also been applied in the study of readability. At the same time, some new technologies in the field of machine learning, such as classification, regression, and sorting, are also used to design new readability evaluation methods, which gives birth to a new measurement method of text readability, that is, the text readability measurement method based on machine learning and complex features [23, 24].

Since the beginning of the twenty-first century, the text readability measurement method based on machine learning and complex features has continuously integrated various rich features and introduced various powerful machine learning frameworks to constantly refresh the performance of the text readability measurement model, which is still developing [24].

With the explosive growth of big data and the emergence of deep learning, a new measurement method has been introduced into the measurement of text readability. The text readability measurement method based on deep learning shows great advantages in measuring the accuracy and automation of text readability [25]. Therefore, this method is a new research trend of text readability measurement methods recently.

However, the research on the measurement of text readability mainly faces several challenges. Firstly, the traditional readability measurement, including readability formula method and measurement method based on artificial intelligence, heavily depends on the extraction of expert artificial features, which seriously lags behind the automation of readability measurement. In the era of big data, how to liberate a large number of labor forces and automatically extract features is a research difficulty. Secondly, with the development of natural language processing and machine learning technology, there are many manually extractable features (e.g., semantic and syntactic structure) that affect the difficulty of text reading. It is more and more difficult to manually extract new features. How to more comprehensively represent the features of text without introducing redundant features is also a difficulty. Thirdly, the measurement of text readability is oriented to different language learners, such as native English (L1) learners and nonnative English (L2) learners. However, the existing model method is difficult to use the same model method to measure the text reading difficulty of L1 and L2. A method that achieves good performance in L1 text readability does not necessarily have the same performance in L2 text.

3. English Text Readability Measurement Based on Convolutional Neural Network

With the rapid development of information technology, it is an era of knowledge explosion and tons of data growth. Finding text materials suitable for the required reading difficulty level in a large number of texts is a very time- and labor-consuming task for readers, which virtually increases their burden. Therefore, effectively measuring the readability of the text and providing readers with intuitive selection criteria will directly affect readers' reading efficiency, which is very necessary. Starting from the various challenges and difficulties faced by the current text readability measurement, this paper regards the text readability measurement task as a classification task and proposes a hybrid network model to measure English text readability based on convolutional neural network (CNN).

The traditional text readability measurement method has some fatal pain points. In the research process of text readability measurement, feature selection excessively

depends on human experts, which limits the development of text readability measurement. There are a wide variety and a large number of existing features that measure text readability. It is more and more difficult to extract new features manually to improve the readability measurement performance, and even introduce redundant and irrelevant features to affect the readability measurement performance.

3.1. CNN Model Introduction. CNN [26] and long short-term memory network (LSTM) [27] are mature and successful deep learning models in the field of natural language processing. Now, these two network models are still the deep learning models that researchers focus on in various natural language related tasks. It is generally believed that CNN is good at capturing local features of language, while LSTM is good at processing sequence data and capturing long-distance dependent information.

In recent years, in order to integrate the advantages of CNN and LSTM, many studies have proposed a hybrid network model based on CNN and LSTM to solve the tasks related to natural language processing.

In order to capture the context information and local features of text, Peng et al. [28] used BiLSTM-DCNN hybrid network model to achieve good performance in text classification task. Fu et al. [29] used CNN-BiLSTM hybrid network model for beautiful sentence recognition. Through experimental comparison with CNN and BiLSTM networks, the results show that the hybrid network model can achieve higher accuracy. Hao et al. [2] also used CNN-BiLSTM hybrid network model to solve the task of Chinese text readability measurement and achieved good performance. For the task of measuring English text readability in this paper, we also use transfer learning and adopt the hybrid network model of CNN and BiLSTM to solve our research problem. CNN is good at extracting local features such as phrases, while BiLSTM can extract text context information and long-distance dependence information. The purpose of constructing this model is to make use of these two advantages [30, 31].

3.2. CNN Model. The structure of CNN model for English text readability measurement is shown in Figure 1.

3.2.1. Word Vector Query Layer. The first layer is the word vector query layer, which is used to mathematically symbolize the natural language sequence to be processed; that is, each given word is projected into the word vector space to facilitate further processing in subsequent layers. The input to this layer is a series of words:

$$\text{Input}_{\text{layer1}} = [w_1, w_2, \dots, w_M]. \quad (1)$$

The output of the query layer is the distributed vector representation of the words queried from GloVe word vector:

$$\bar{x} = [x_1, x_2, \dots, x_n], \quad (2)$$

where $x_i \in R^d$ and n is the length of the sequence.

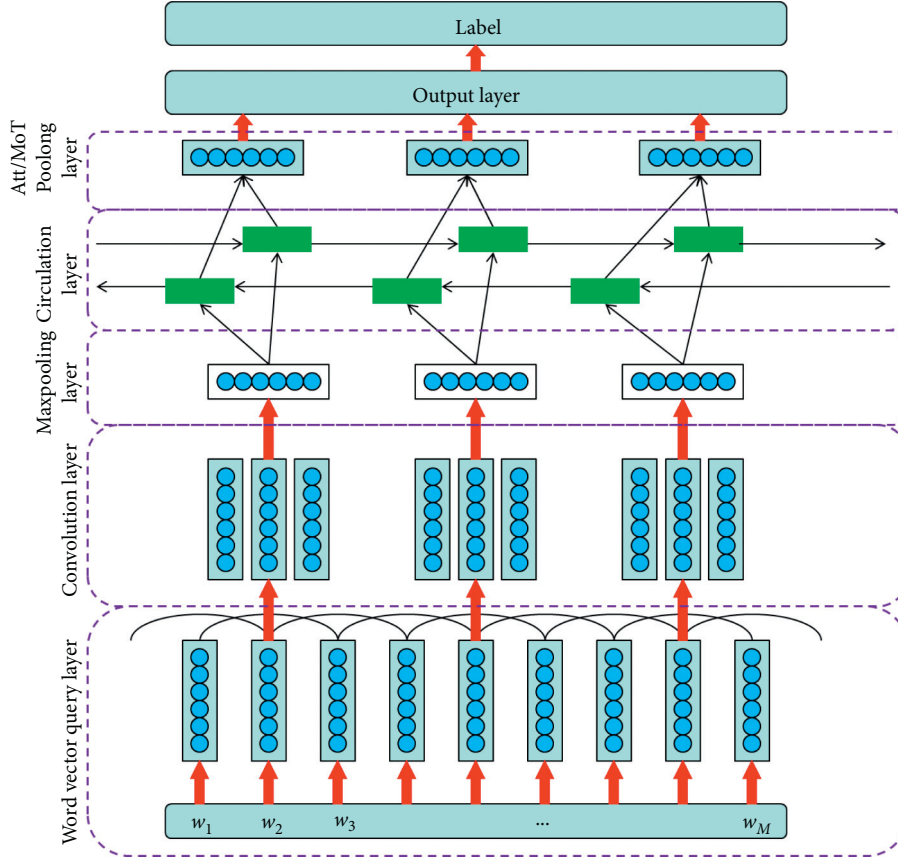


FIGURE 1: CNN model structure.

3.2.2. Convolution Layer and Max-Pooling Layer. Once the word vector representation \bar{x} of the input sequence is queried, in order to more comprehensively extract local features from the sequence, the convolution layer will use multiple filters of different sizes to continuously perform convolution operation on the word vector sequence \bar{x} by sliding.

If the filter size of the convolution layer is k , the filter can be expressed as a matrix:

$$m \in R^{k \times d}. \quad (3)$$

In the filter sliding process, for each position i in the sequence, there is a window matrix \bar{w}_i with k consecutive words, expressed as

$$\bar{w}_i = [x_i, x_{i+1}, \dots, x_{i+k-1}]. \quad (4)$$

The filter matrix m is convoluted with the word window matrix \bar{w}_i (k -gram) at each position in an effective way to generate a feature map:

$$c \in R^{L-k+1} \quad (5)$$

The feature mapping of the word window vector \bar{w} at position i can be calculated as

$$c_i = \sigma(\bar{w} \otimes m + b), \quad (6)$$

where \otimes is multiplication, b is bias, and σ is the activation function of sigmoid.

Then, in the convolution layer, the max pooling will be further used for the results of convolution calculation. The max pooling will filter the maximum value in c_i as the feature of the filter corresponding to the i th word. The max pooling can reduce the output parameters of CNN and the risk of overfitting and also reduce the impact of filling 0 when processing input sequences of equal length.

In terms of convolution operation mode, convolution layer is similar to n -gram language model. It is good at extracting local context information in article sequence, so as to improve the performance of the model.

3.2.3. Circulation Layer. After generating the embedding (whether from the convolution layer or directly from the query layer), the loop layer starts processing the input sequence to generate a representation of a given article. Ideally, the representation can encode all the information needed to measure text readability. However, because the text is usually very long and consists of hundreds of word sequences, the vector representation learned by the final state of the loop layer may not be enough for accurate readability measurement.

For this reason, we keep all the intermediate states of the loop layer so that we can track and process the important information of the article. For the circular layer, based on the experimental experience, we choose BiLSTM to extract the long-distance dependence information of text context and sequence.

In order to control the information flow during the processing of the input sequence, the LSTM uses three gates to forget or remember the transmitted information of the sequence. The functions of LSTM are described as follows:

$$\begin{aligned}
i_t &= \sigma(W_i \cdot X_t + U_i \cdot h_{t-1} + b_i), \\
f_t &= \sigma(W_f \cdot X_t + U_f \cdot h_{t-1} + b_f), \\
\tilde{c}_t &= \tanh(W_c \cdot X_t + U_c \cdot h_{t-1} + b_c), \\
c_t &= i_t \circ \tilde{c}_t + f_t \circ c_{t-1}, \\
o_t &= \sigma(W_o \cdot X_t + U_o \cdot h_{t-1} + b_o), \\
h_t &= o_t \circ \tanh(c_t),
\end{aligned} \tag{7}$$

where X_t and h_t are the input and output vectors at time t , respectively; $W_i, W_f, W_c, W_o, U_i, U_f, U_c, U_o$ are weight matrices; b_i, b_f, b_c, b_o are bias vectors; and the symbol \circ represents element-by-element multiplication.

3.2.4. ATT/MoT Pooling Layer. This layer is connected behind the circulating layer and receives the output $H = (h_1, h_2, \dots, h_M)$ of the circulating layer. It is responsible for aggregating the variable length input H into a fixed length vector, so as to facilitate the use of subsequent network layers. There are generally two common methods for this layer, mean over time and attention pooling.

(1) *Mean over time method.* The average time layer receives M vectors with dimension d_r as input and calculates average vectors of the same length. The calculation formula is defined as follows:

$$v = \frac{\sum_{i=1}^M h_i}{M} \tag{8}$$

After the vector is calculated, it is sent to the subsequent network layer for corresponding operation.

(2) *Attention-pooling method.* The average time layer is equivalent to assigning an equal weight $1/M$ to the output H of the cycle layer, and the average time layer can also be replaced by a self-attention mechanism. The self-attention mechanism can learn the importance of the output of each intermediate state of the loop layer to the characterization of the whole document and assign a weight α_i to each output state h_i . The calculation formula is defined as follows:

$$\begin{aligned}
u &= \tanh(W h_i + b), \\
\alpha_i &= \frac{\exp(u^T u_w)}{\sum_i \exp(u^T u_w)}, \\
v &= \sum_i \alpha_i h_i,
\end{aligned} \tag{9}$$

where h_i represents the output of the intermediate state of the loop layer and u_w represents the vector representing the text context information. This vector is an initialization vector and will be automatically learned in the backpropagation.

3.2.5. Softmax Layer. The final representation vector v of the text is obtained from the previous pooling layer and then sent to the softmax layer for classification. In this CNN model, cross entropy is selected as the loss function.

4. Case Study

4.1. Data Set. The existing gold data sets for English text readability measurement are Weekly Reader data set and WeeBit data set. In particular, WeeBit data set is one of the most popular data sets in text readability measurement tasks, with the largest amount of data and the most standard readability label [32].

WeeBit data set consists of two parts of data. The first part is Weekly Reader corpus, which is also one of the popular gold data sets in English text readability measurement tasks. The corpus comes from Weekly Reader (<https://www.weeklyreader.com>), an American educational news magazine. The texts in the magazine are compiled by educational experts according to the readers' age, and their age groups are designated as the reading difficulty level of the corresponding texts. The text content of the magazine is mainly applicable to the reading objects at ages 7-8, 8-9, 9-10, and 10-12. Another part of the data comes from the BBC Bitesize website, which provides readers of different ages with articles of corresponding difficulty. The WeeBit corpus uses text data corresponding to reading difficulty from two age groups on the BBC Bitesize website, which are 11-14 years old and 14-16 years old, respectively. These two parts of data are combined to form WeeBit corpus.

Because these two data sets are popular and authoritative in the field of English text readability measurement and in order to better compare our experimental results with existing methods, we use these two data sets. The details of these two data sets are shown in Table 1.

4.2. Evaluating Indicator. This paper uses the two most commonly used evaluation indicators in text readability measurement tasks, accuracy and Pearson correlation.

4.2.1. Accuracy. In this paper, ACC is used to express the accuracy. We suppose that there are two types of original samples: P positive samples in total, marked as 1; N negative samples, marked as 0. After classification, TP samples with category 1 are correctly determined as 1 by the model, and FN samples with category 1 are determined as 0 by the model. Obviously, $P = TP + FN$. FP samples with category 0 are correctly determined as 1 by the model, and TN samples with category 0 are determined as 0 by the model. Obviously, $N = FP + TN$. Then, ACC can be defined as follows:

TABLE 1: The details of Weekly Reader corpus and WeeBit corpus.

	Reading level	Applicable age	Number of chapters	Average number of sentences per text
Weekly Reader corpus	Level 2	7-8	633	23.45
	Level 3	8-9	795	23.22
	Level 4	9-10	805	29.17
	Senior	10-12	1316	31.22
WeeBit corpus	Level 2	7-8	641	23.01
	Level 3	8-9	791	23.45
	Level 4	9-10	822	29.23
	KS3	11-14	652	22.11
	GCSE	14-16	3600	28.22

$$ACC = \frac{TP + TN}{P + N}. \quad (10)$$

Accuracy (ACC) reflects the classifier’s ability to classify the whole sample, that is, the ability to classify positive samples as positive and negative samples as negative.

4.2.2. *Pearson correlation.* In this paper, PCC is used to express the Pearson correlation. PCC is defined as the quotient of covariance and standard deviation between two sequence variables, which is as follows:

$$PCC(X, Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}}. \quad (11)$$

PCC can describe the correlation between two sequences X and Y , and the value range of PCC is $[-1, 1]$. When $PCC > 0$, X and Y are positively correlated. When $PCC < 0$, X and Y are negatively correlated. When $PCC = 0$, the two variables are not related. Generally, the greater the absolute value of PCC, the stronger the correlation between variables; that is, the closer the PCC to 1 or -1 , the stronger the correlation. The closer the PCC to 0, the weaker the correlation. In the process of text readability measurement, the value range in Table 2 is usually used to judge the correlation strength between the two sequences.

4.3. *Experimental Environment and Super Parameter Settings.* The experimental environment and super parameter settings are shown Table 3. The laboratory is completed under Ubuntu system (Python 3.5 version), and other environment parameters are shown in Table 3.

In the setting of super parameters, it is unrealistic to find the learning rate that can make the convergence speed of loss function moderate and find the global optimal solution based on personal experience or multiple experiments, so the learning rate of the network model is set as the initial value, that is, the dynamic learning rate of 0.001, and the corresponding parameters are automatically updated during model training. A learning rate which is more suitable for the model can be found. In the convolution layer of the first level network of the hybrid network model, in order for the model to capture local information more comprehensively, such as phrase information with different lengths, we set the size of the convolution kernel to 3, 4, and 5; extract the

TABLE 2: PCC value range and its correlation strength.

PCC value	Correlation strength
0-0.2	Very weak correlation or no correlation
0.2-0.4	Weak correlation
0.4-0.6	Moderate correlation
0.6-0.8	Strong correlation
0.8-1	Extremely strong correlation

TABLE 3: Experimental environment parameters.

Name	Parameter
Memory	15.6 G
Graphics	GeForce GTX 1080 Ti/PCLe/SSE2
Processor	Intel Core™ i7-8700 CPU @ 3.7 GHz x 12

corresponding features, respectively; and splice them into the total features.

Other hyperparameter settings are shown in Table 4.

4.4. Analysis of Experimental Results

4.4.1. *Comparison with CNN and LSTM Related Models.* As shown in the experimental results in Table 5, we conducted experiments on various models related to CNN and LSTM on the gold standard data set WeeBit and compared them with the hybrid network model proposed in this paper.

Firstly, as we know, compared with the long short-term memory network, which only considers the following information: the bidirectional long short-term memory network can extract effective long-distance dependence and other information because it considers the context information. Therefore, CNN and BiLSTM are selected in the hybrid network model. Secondly, in theory, long short-term memory networks (including LSTM and BiLSTM) should be better at dealing with the task of sequence data input than convolutional neural network (CNN), but from the experimental results, the performance of long short-term memory network is slightly inferior to convolutional neural network model. The reason for this result may be that our network model takes the whole text sequence composed of word sequence as the input, and the length of the sequence is uncertain and long, which limits the performance of LSTM to a certain extent. Finally, in the hybrid network model, the final text representation can be calculated directly using the final state of LSTM or BiLSTM instead of the output of the

TABLE 4: Hyperparameter settings.

Hyperparameter	Introduction	Value
learning.rate	Initial value of learning rate	0.001
embedding.size	Word vector dimension	100
filter.size	Convolution kernel size	3,4,5
num.filter	Number of convolution kernels	200
Dropout	Dropout probability size	0.5
l2.reg.lambda	Size of L2 regularized lambda	0.0001
lstm.hidden	LSTM hidden layer size	100
batch.size	Batch size	100
max.length	Length of sequence	1538

TABLE 5: Comparison with CNN and LSTM related models.

Model	Accuracy	Pearson correlation coefficient
CNN	0.801	0.840
LSTM	0.711	0.744
BiLSTM	0.719	0.836
CNN-BiLSTM	0.831	0.892
CNN-BiLSTM-MoT	0.877	0.921
CNN-BiLSTM-ATT	0.886	0.938

TABLE 6: Comparison with existing traditional methods (on WeeBit data set).

Model	Accuracy	Pearson correlation coefficient
Model 1 [33]	0.929	—
Model 2 [34]	0.811	0.902
The proposed model	0.891	0.932

TABLE 7: Comparison with existing traditional methods (on Weekly Reader data set).

Model	Accuracy	Pearson correlation coefficient
Model 3 [30]	0.732	—
Model 4 [31]	0.628	—
Model 1 [33]	0.911	—
The proposed model	0.775	0.836

intermediate state of the cyclic network. However, experiments show that it is better to retain the output of the intermediate state and connect the pooling layer. Moreover, we also need to properly consider the selection of the pooling layer connected after LSTM output. Considering the use of the attention mechanism layer in the pooling layer will get the best model effect and can achieve an accuracy (ACC) of 0.886 and a Pearson correlation coefficient (PCC) of 0.938 on the WeeBit data set.

4.4.2. Comparison with the Existing Traditional Methods.

As shown in Tables 6 and 7, we use the hybrid network model to do empirical research on WeeBit data set and Weekly Reader data set, respectively, and compare the experimental results with the results of existing model methods on the corresponding data set.

The experimental results show that the accuracy (ACC) of the proposed hybrid network model is 0.891 and the Pearson correlation coefficient is 0.932 on the WeeBit data set, while the

accuracy (ACC) of 0.775 and Pearson correlation coefficient (PCC) of 0.836 are obtained on the Weekly Reader data set. It can be seen from the table that under the two measurement indices of accuracy and Pearson correlation, the performance of this hybrid network model is better than most traditional methods, but it is slightly inferior to the best model methods. In general, the hybrid network model has achieved competitive performance compared with traditional methods. In particular, it can automatically extract text readability related features, completely replace labor, liberate labor, and greatly improve the practicability of the model method in the task of text readability measurement.

5. Conclusions

Traditional text readability measurement methods have some fatal pain points. In the research process of text readability measurement, feature selection excessively depends on human experts, which limits the development of text readability measurement. There are a wide variety and a

large number of existing features that measure text readability. It is more and more difficult to extract new features manually to improve the readability measurement performance, and even introduce redundant and irrelevant features to affect the readability measurement performance. To solve these problems, this paper proposes a hybrid network model for text readability measurement based on convolutional neural network, makes an empirical study on this method, and evaluates the performance of the model.

The proposed hybrid network model based on convolutional neural network has limitations or deficiencies in measuring text readability. Firstly, the hybrid network model regards the whole document as a sequence composed of one word. Due to the different length of the text, it will be filled with 0 in the process of processing the equal length input, which will introduce a lot of redundant information into the sequence features extracted by the network model, which will affect the performance of the model. Secondly, because the whole document is directly used as an input sequence, the factors that can affect the readability of the text contained in the sentence related information in the text (such as the logical structure and syntactic relationship between sentences) will be lost. Finally, because the network model takes the word sequence of the whole document as the input, the text will be relatively long theoretically, ranging from hundreds of words to thousands of words. In the process of processing such a long sequence, the gradient will disappear. With the continuous growth of the sequence length, some context information will be lost after long-distance information transmission. This limits the performance of long short-term memory networks. Future research will focus on how to overcome the limitations of hybrid network model and the construction of improved hybrid network model, such as hierarchical hybrid network model.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] J. Galliussi, L. Perondi, G. Chia, W. Gerbino, and P. Bernardis, "Inter-letter spacing, inter-word spacing, and font with dyslexia-friendly features: testing text readability in people with and without dyslexia," *Annals of Dyslexia*, vol. 70, no. 1, pp. 141–152, 2020.
- [2] L. Hao, L. Si, J. Zhao, Z. Bao, and X. Bai, "Chinese Teaching Material Readability Assessment with Contextual information," in *Proceedings of the International Conference on Asian Language Processing*, IEEE, Singapore, December 2017.
- [3] H. Mohammadi and S. H. Khasteh, "A machine learning approach to Persian text readability assessment using a crowdsourced dataset," in *Proceedings of the 28th Iranian conference on electrical engineering (ICEE)*, Tabriz Iran, August 2020.
- [4] O. Kodym and M. Hradiš, "TG(2): text-guided transformer GAN for restoring document readability and perceived quality," *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 9, pp. 1–14, 2021.
- [5] S. Nahatame and P. Effort, "Text Readability In second language reading: a computational and eye-tracking investigation," *Language learning*, vol. 71, no. 4, pp. 1004–1043, 2021.
- [6] A. Ac, C. Gs, and B. Rc, "An online multi-source summarization algorithm for text readability in topic-based search," *Computer Speech & Language*, vol. 66, Article ID 101143, 2020.
- [7] X. Chen and D. Meurers, "Word frequency and readability: predicting the textlevel readability with a lexiclevel attribute," *Journal of Research in Reading*, vol. 41, no. 3, pp. 486–510, 2018.
- [8] M. Aha, I. Ehtesham, T. Konstantinos et al., "Motor signs in Alzheimer's disease and vascular dementia: detection through natural language processing, co-morbid features and relationship to adverse outcomes - ScienceDirect," *Experimental Gerontology*, vol. 146, 2021.
- [9] E. C. A. Carreón, H. Espaa, H. Nonaka, and T. Hiraoka, "Differences in Chinese and Western tourists faced with Japanese hospitality: a natural language processing approach," *Information Technology & Tourism*, vol. 23, no. 2, pp. 381–438, 2021.
- [10] S. Vajjala and D. Meurers, *Readability-based sentence ranking for evaluating text simplification*, Iowa State University, Ames, Iowa, 2016.
- [11] A. A. Al-Ajlan, H. S. Al-Khalifa, and A. Al-Salman, "Towards the Development of an Automatic Readability Measurements for Arabic language," in *Proceedings of the International Conference on Digital Information Management*, IEEE, London, UK, November 2008.
- [12] D. Alejos, P. Tregubenko, and A. Kumar, "Preservation of life is not easy: readability text analysis of patient information on fertility preservation options," *Clinical Lymphoma, Myeloma & Leukemia*, vol. 19, 2019.
- [13] H. C. Tseng, B. Chin, T. H. Chang, and Y. T. Sung, "Integrating LSA-based hierarchical conceptual space and machine learning methods for leveling the readability of domain-specific texts," *Natural Language Engineering*, vol. 25, no. 3, pp. 1–31, 2019.
- [14] L. La, N. Wang, and D. P. Zhou, "Improving reading comprehension step by step using Online-Boost text readability classification system," *Neural Computing & Applications*, vol. 26, no. 4, pp. 929–939, 2015.
- [15] M. Gattullo, A. E. Uva, M. Fiorentino, and G. Monn, "Effect of text outline and contrast polarity on AR text readability in industrial lighting," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 5, pp. 638–651, 2015.
- [16] S. T. Yao, L. C. Wei, S. B. Dyson, K. Chang, and Y. Chen, "Leveling L2 texts through readability: combining multilevel linguistic features with the CEFR," *The Modern Language Journal*, vol. 99, no. 2, pp. 371–391, 2015.
- [17] B. J. Esfahani, A. Faron, K. S. Roth, P. P. Grimminger, and J. C. Luers, "[Systematic readability analysis of medical texts on websites of German university clinics for general and abdominal surgery]," *Zentralbl Chir*, vol. 141, no. 6, 2016.
- [18] A. Schlapbach, F. Wettstein, and H. Bunke, "Estimating the Readability of Handwritten Text - a Support Vector Regression Based approach," in *Proceedings of the International Conference on Pattern Recognition*, IEEE, Tampa, FL, USA, December 2008.
- [19] J. M. O. Amp and R. A. R. King, "A matter of significance: can sampling error invalidate cloze estimates of text readability? [J]," *Language Assessment Quarterly*, vol. 7, no. 4, pp. 303–316, 2010.

- [20] Y. Gecit, "The evaluation of high school geography 9 and high school geography 11 text books with some formulas of readability," *Educational ences: Theory and Practice*, vol. 10, no. 4, pp. 2205–2220, 2010.
- [21] R. G. Benjamin, "Reconstructing readability: recent developments and recommendations in the analysis of text difficulty," *Educational Psychology Review*, vol. 24, no. 1, pp. 63–88, 2012.
- [22] U.-V. Marti and H. Bunke, "Using a statistical language model to improve the performance of an hmm-based cursive handwriting recognition system," *Series in Machine Perception and Artificial Intelligence*, vol. 15, no. 1, pp. 65–90, 2001.
- [23] D. Jones, E. Gibson, W. Shen et al., "Measuring human readability of machine generated text: three case studies in speech recognition and machine translation," in *Proceedings of the (ICASSP '05)IEEE International Conference on Acoustics Speech and Signal Processing 2005*, IEEE, Philadelphia PA USA, March 2005.
- [24] R. Shams and R. E. Mercer, "Classifying spam emails using text and Features," in *Proceedings of the IEEE International Conference on Data Mining*, pp. 657–666, IEEE, Dallas TX USA, December 2013.
- [25] Y. X. Sun, K. Y. Chen, L. Sun, and C. Hu, "Attention-based deep learning model for text readability evaluation," in *Proceedings of the International joint conference on neural networks (IJCNN) held as part of the IEEE world congress on computational intelligence (IEEE WCCI)*, IEEE, Glasgow UK, July 2020.
- [26] A. N. Gorban, E. M. Mirkes, and I. Y. Tugin, "How deep should be the depth of convolutional neural networks: a backyard dog case study," *Cognitive Computation*, vol. 12, no. 1, pp. 388–397, 2020.
- [27] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," *Computer Science*, vol. 20, pp. 338–342, 2014.
- [28] Z. Peng, Z. Qi, S. Zheng, J. Xu, H. Bao, and B. Xu, "Text classification improved by integrating bidirectional LSTM with two-dimensional max pooling," in *Proceedings of the 26th International Conference on Computational Linguistics*, Technical Papers, Osaka Japan, December 2016.
- [29] R. Fu, D. Wang, S. Wang, G. Hu, and T. Liu, "Elegart sentence recognition for automated essay scoring," *Journal of Chinese Information Processing*, vol. 32, no. 6, p. 10, 2018.
- [30] C. Wightingon, S. Stewart, B. Davis, B. Barrett, B. Price, and S. Cohen, "Data augmentation for recognition of handwritten words and lines using a CNN-LSTM network," in *Proceedings of the 2017 14th IAPR international conference on document analysis and recognition (ICDAR)*, Kyoto Japan, November 2017.
- [31] N. Sadeghzadehyazdi, T. Batabyal, and S. T. Acton, "Modeling spatiotemporal patterns of gait anomaly with a CNN-LSTM deep neural network," *Expert Systems with Applications*, vol. 185, Article ID 115582, 2021.
- [32] S. Vajjala and D. Meurers, "On Improving the Accuracy of Readability Classification Using Insights from Second Language Acquisition," in *Proceedings of the Workshop on Innovative Use of Nlp for Building Educational Applications*, Association for Computational Linguistics, Montreal Canada, June 2012.
- [33] J. Hancke, S. Vajjala, and D. Meurers, *Readability classification for German using lexical, syntactic, and morphological features*, Springer international publishing, Manhattan, New York City, 2014.
- [34] M. Xia, E. Kochmar, and T. Briscoe, "Text readability assessment for second language learners," in *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications*, CA, USA, June 2016.

Research Article

Coal Mine Safety Evaluation Based on Machine Learning: A BP Neural Network Model

Guangxing Bai ¹ and Tianlong Xu²

¹College of Safety Science and Engineering, Xian University of Science and Technology, Shaanxi, Xian 710054, China

²College of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 611730, Sichuan, China

Correspondence should be addressed to Guangxing Bai; baiguangxing@xust.edu.cn

Received 9 January 2022; Revised 20 January 2022; Accepted 9 February 2022; Published 14 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Guangxing Bai and Tianlong Xu. 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.

As the core of artificial intelligence, machine learning has strong application advantages in multi-criteria intelligent evaluation and decision-making. The level of sustainable development is of great significance to the safety evaluation of coal mining enterprises. BP neural network is a classical algorithm model in machine learning. In this paper, the BP neural network is applied to the sustainable development level decision-making and safety evaluation of coal mining enterprises. Based on the analysis of the evaluation method for sustainable development of coal enterprises, the evaluation index system of sustainable development of coal enterprises is established, and a multi-layer forward neural network model based on error backpropagation algorithm is constructed. Based on the system theory of man, machine, environment, and management, and taking the four single elements and the whole system in a coal mine as the research object, this paper systematically analyzes and studies the evaluation and continuous improvement of coal mine intrinsic safety. The BP neural network evaluation model is used to analyze and study the intrinsic safety of coal mines, the shortcomings of the intrinsic safety construction of coal mines are found, and then improvement measures are put forward to effectively promote the safe production of coal mines and finally realize the intrinsic safety goal of the coal mine.

1. Introduction

Coal will still be the main energy source for a long time. At present, the rapid growth of the economy puts forward higher requirements for the development of the coal industry [1–6]. Therefore, we must strengthen safety production and ensure the sustainable, stable, and healthy development of the coal industry. However, the coal industry is a high-risk industry. High gas and gas outburst coal mines account for about half of China's coal mines. Coal mine safety is the top priority of the whole industrial safety production work. Coal mining enterprises have the characteristics of many personnel, scattered operations, many equipment and facilities, wide distribution, bad natural conditions, many unsafe factors, complex working environments, and difficult managements. The workplace is

constantly changing [7]. The risk factors of natural disasters and production accidents always affect and restrict the safe production of coal mines. On the other hand, in recent years, with the continuous changes in the internal and external environment faced by coal enterprises and the continuous deepening of the reform of large and medium-sized state-owned enterprises, the operating conditions of coal enterprises have fluctuated. On the whole, they are in the process of continuous adaptation and re-adaptation, organization and reorganization, and innovation and re-innovation [8–11]. Coal resources are nonrenewable resources. Coal mining is bound to be restricted by the remaining reserves in the mining area, and coal enterprises will face resource depletion sooner or later. Therefore, the problem of sustainable development of coal enterprises is becoming increasingly prominent. Therefore, it is very necessary to

construct a coal mine safety evaluation model based on the research on the evaluation of the sustainable development level of coal enterprises.

Academic circles and decision-making departments at home and abroad have made a lot of exploration, especially in the evaluation of the sustainable development level of coal enterprises. Machine learning can be regarded as a task. The goal of this task is to let machines (computers in a broad sense) acquire human-like intelligence through learning. A neural network is a method to realize machine learning tasks. Talking about the neural networks in the field of machine learning generally refers to “neural network learning.” It is a network structure composed of many simple units. This network structure is similar to the biological nervous system, which is used to simulate the interaction between organisms and the natural environment. An artificial neural network (ANN) is an information processing system imitating human brain model [12,13]. It has good abilities of self-learning, self-adaptation, associative memory, parallel processing, and nonlinear transformation.

How to effectively curb the occurrence of major mining accidents is the biggest problem to be solved in China’s coal mine production. Coal mine safety theory is put forward in this environment. The coal mine underground is a complex and changeable man-machine environmental system. This paper attempts to evaluate the sustainable development level of coal enterprises by establishing a multi-layer forward neural network model based on the error back propagation algorithm (BP algorithm). It can avoid complex mathematical derivation and ensure stable results in the case of sample defect and parameter drift. It can also effectively avoid the classical sustainable development evaluation methods, such as the analytic hierarchy process [14–16], fuzzy mathematics [17–22], and principal component analysis [23,24] and cannot avoid the role of people’s experience and knowledge and the personal subjective intention of decision-makers, which is of great benefit to solve the overall decision-making planning of coal enterprises. This paper will use the system theory to take the coal mine man-machine-environment-management system as the research object, establish the coal mine intrinsic safety evaluation system and evaluation model, comprehensively construct the coal mine intrinsic safety system through the specific and in-depth analysis of various factors of man-machine-environment-management, provide the basis for coal mine safety production and management, and improve the safety production level of the coal mine industry.

2. BP Neural Network Model for Sustainable Development of Coal Enterprises

2.1. Model Building. According to the meaning of sustainable development of coal enterprises and the principle of index system design, combined with the existing achievements and the research on the specific situation of coal enterprises, an index system including 5 criteria layers and 17 specific indicators is constructed, as shown in Figure 1. According to the evaluated problems, combined with the

multi-layer forward neural network model based on the error backpropagation algorithm (BP algorithm), the neural network model for sustainable development of coal enterprises is established, as shown in Figure 2.

The model is divided into two modules: the former is the normalization module, and the latter is the BP neural network (BPNN) module [25–28]. The BPNN module in the above model adopts a three-layer BP neural network, including an input layer, a hidden layer, and an output layer P. The input of a neural network is required to be in [0 and 1], so the original data of each evaluation index shall be normalized before network learning and training. The specific normalization rules are shown in Table 1. In this way, the network input value corresponding to each evaluation index in the sample can be determined by normalization.

2.2. Network Training and Learning. The original data are sent to the normalization module after preprocessing. The normalization module will normalize the input data according to the rules in Table 1 to obtain 17 normalized values, and then input the normalized values into the BPNN module. According to the above analysis, the number of fuzzy neurons in the input layer of the BPNN module is 17; that is, the input signals x_1, x_2, \dots, x_{17} correspond to 17 normalized values; the number of output neurons is 1, i.e., output o , which corresponds to the sustainable development level of coal enterprises. The number k of neurons in the hidden layer was adjusted by the learning process to 35.

The learning process of the BP neural network is also the process of network parameter correction. The network learning system adopts the method with teachers, and the correction of network parameters adopts the gradient method. It is assumed that there are n system sample data: $\{\hat{O}_a, O_a\}$, $a = 1, 2, \dots, n$. Here, the subscript a represents the sample serial number, \hat{O}_a is the sample output, and O_a is the actual output. x_{ia} is input variable, $i = 1, 2, \dots, 17$. The input variable will be assigned to the m -th neuron of the hidden layer as its input according to the following formula:

$$x'_m = \sum_{i=1}^{17} w_{im} x_{ia}, \quad (1)$$

where w_{im} is the weight of the input layer neuron i and the hidden layer neuron m .

The most commonly used transfer function of BP neuron is the sigmoid function:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2)$$

According to the sigmoid function, it is obtained that the function of the output O'_m of the hidden layer neuron m with respect to the input x'_m is

$$O'_m = \frac{1}{1 + e^{-x'_m}}. \quad (3)$$

Similarly, the input and output of each unit of the output layer can also be obtained, which will not be described in detail here.

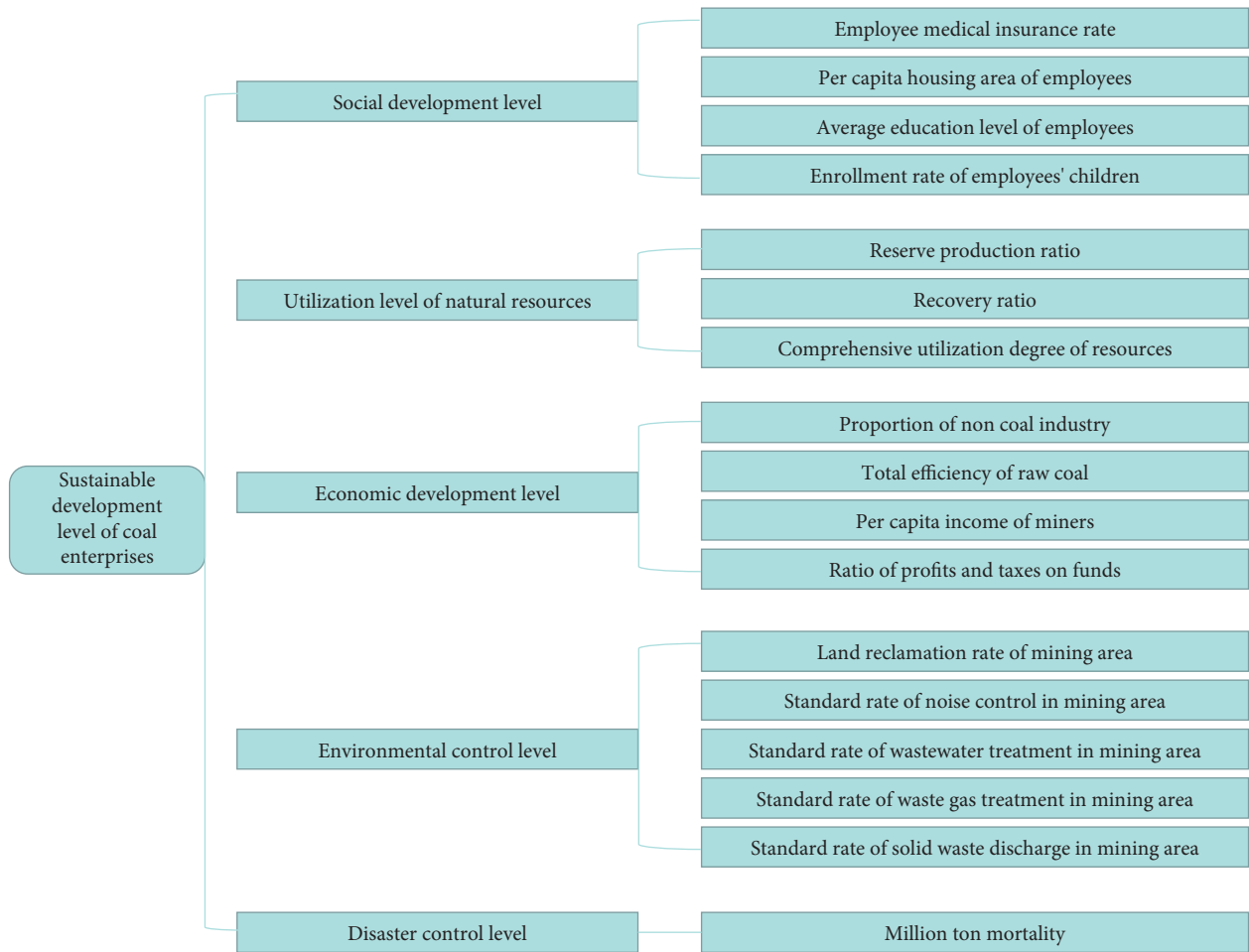


FIGURE 1: Evaluation index system of sustainable development of coal enterprises.

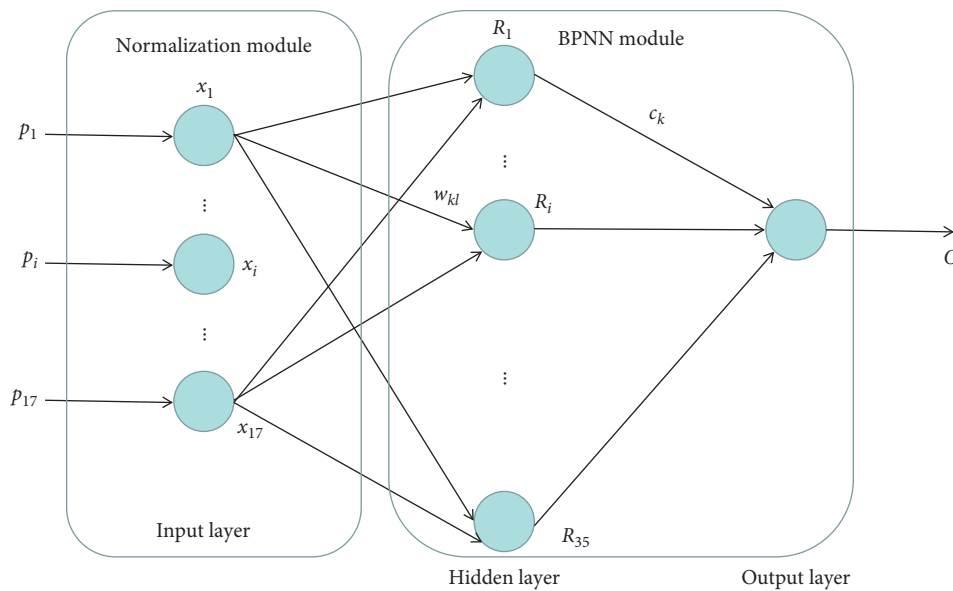


FIGURE 2: BP neural network model.

Through a certain number of network training processes, it is actually to modify the network parameter to determine the most appropriate weight, so as to minimize the residual

error between the actual output O_a and the sample output \hat{O}_a obtained by forward operation according to equations (1) and (3) for all n sample inputs. The residual error is as follows:

TABLE 1: Normalization rules of neural network input data.

Evaluation index	Network input					
	Range	Value	Range	Value	Range	Value
Employee medical insurance rate (%)	≤ 10	0.1	(10, 100)	Y/100	100	1
Per capita housing area of employees (m ² /person)	$= 0$	0	(0, 20)	Y/20	≥ 20	1
Average education level of employees	Below medium	0.3–0.5	Fairly good	0.7	Good	0.9
Enrollment rate of employees' children (%)	≤ 10	0.1	(10, 100)	(Y-80)/(120-80)	100	1
Reserve production ratio (%)	≤ 80	0.1	(80, 120)	Y/100	≥ 120	1
Recovery ratio (%)	≤ 10	0.1	(10, 100)	Y/100	100	1
Comprehensive utilization degree of resources	Below medium	0.3–0.5	Fairly good	0.7	Good	0.9
Proportion of noncoal industry (%)	≤ 10	0.1	(10, 100)	Y/100	100	1
Total efficiency of raw coal (t/person)	$= 0$	0	(0, 4)	Y/4	≥ 4	1
Per capita income of miners (yuan/month)	≤ 1000	0.1	(1000, 10000)	Y/10000	≥ 10000	1
Ratio of profits and taxes on funds (%)	$= 0$	0	(0, 20)	Y/20	≥ 20	1
Land reclamation rate of mining area (%)	≤ 10	0.1	(10, 100)	Y/100	100	1
Standard rate of noise control in mining area (%)	≤ 10	0.1	(10, 100)	Y/100	100	1
Standard rate of wastewater treatment in mining area (%)	≤ 10	0.1	(10, 100)	Y/100	100	1
Standard rate of waste gas treatment in mining area (%)	≤ 10	0.1	(10, 100)	Y/100	100	1
Standard rate of solid waste discharge in mining area (%)	≤ 10	0.1	(10, 100)	Y/100	100	1
Million ton mortality (%)	$= 0$	1	(0, 1)	(1-Y)/1	≥ 1	0

$$E = \frac{1}{2} \sum_{a=1}^n (O_a - \hat{O}_a)^2. \quad (4)$$

The correction of weight and threshold is realized by the gradient method of the back propagation algorithm. t represents the time of iterative correction, and b_k and b_o represent the neuron thresholds of the hidden layer and output layer, respectively, then the parameter correction rule of the BP neural network is

- (1) The connection weight from the input layer to the hidden layer is

$$w_{ki}(t+1) = w_{ki}(t) - \eta \frac{\partial E}{\partial b_k}, \quad (5)$$

where $i = 1, 2, \dots, 17$; $k = 1, 2, \dots, 35$; w_{ki} is the connection weight from the input node x_i to the hidden layer node R_k ; and η is the learning rate.

- (2) Hidden layer neuron threshold is

$$b_k(t+1) = b_k(t) - \eta' \frac{\partial E}{\partial w_{ki}(t)}, \quad (6)$$

where $k = 1, 2, \dots, 35$ and η' is the learning rate.

- (3) Connection weight from hidden layer to output layer is

$$c_k(t+1) = c_k(t) - \eta'' \frac{\partial E}{\partial c_k}, \quad (7)$$

where $k = 1, 2, \dots, 35$; c_k is the weight from the rule layer node R_k to the output layer node O ; and η'' is the learning rate.

- (4) Output layer neuron threshold is

$$b_o(t+1) = b_o(t) - \eta''' \frac{\partial E}{\partial b_o}, \quad (8)$$

where η''' is the learning rate.

After training and learning, the evaluation network can output the evaluation value to measure the level of sustainable development, which ranges from [0,1]. In order to clarify the sustainable development level of coal enterprises, the sustainable development status is divided into four levels: the first level is sustainable development, and the score range is $0.85 < \beta \leq 1$; the second level is primary sustainable development, and the score range is $0.70 < \beta \leq 0.85$; the third level is the transition from traditional development to sustainable development, and the score range is $0.50 < \beta \leq 0.70$; and the fourth level is traditional development, and the score range is $0 < \beta \leq 0.50$. In this way, the sustainable development level of the enterprise can be clearly obtained from the network output value. In each evaluation work, no matter whether the evaluation result is recognized by experts or not, it can be used as a new learning sample to make the BP neural network evaluation system learn and improve continuously, so as to make it make a more accurate evaluation.

3. Comprehensive Evaluation Model of Coal Mine Safety

This study establishes the safety evaluation index system of each element from the four elements of man, machine, environment, and management [29,30]. The construction of a coal mine safety evaluation model needs to organically combine the four elements of man, machine, environment, and management. Therefore, man, machine, environment, and management can be regarded as four primary indicators. Among the secondary indicators, human intrinsic safety indicators are divided into physical status, psychological status, safety education status, and safety technology status. Equipment safety indicators are divided into equipment reliability and production system factors. Environmental essential indicators can be divided into two

categories: geological environment and working environment. The management indicators are divided into personal injury and loss, training and education, intrinsic safety management system establishment, safety measures, basic management, emergency rescue, and enterprise safety culture.

Combining the four elements organically, a coal mine safety evaluation classification model is constructed, which can be divided into four primary evaluation indexes and 14 secondary evaluation indexes, as shown in Figure 3. The scoring standard for coal mine safety evaluation in Table 2 is established with reference to the national guiding principles of intrinsic safety.

4. Fuzzy Evaluation of Coal Mine Safety Based on BP Neural Network

Fuzzy neural network (FNN) is a new and better system combining neural networks and fuzzy logic systems [31–33]. The system not only has the advantages of a neural network, that is, it has the function of self-organizing and adaptive learning, but also makes up for the deficiency of a neural network, that is, it can directly deal with structured knowledge. The weights without clear network meaning in the traditional neural network give the physical meaning of the rule parameters in the fuzzy system, which is convenient to use the rule parameters to study things.

4.1. Fuzzy Neural Network Learning Algorithm. It is assumed that n and m are the numbers of input units and hidden units, respectively. $X = (x_1, x_2, \dots, x_n)$ is the input layer input of the fuzzy system. After fuzzy processing of membership function, $R = (r_1, r_2, \dots, r_n)$ is obtained, which is the input vector of the neural network. $Z = (z_1, z_2, \dots, z_n)$ is the hidden layer output vector and $Y = (y_1, y_2, \dots, y_n)$ is the system output vector.

$W_j = (w_1, w_2, \dots, w_{jn})$ is the weight vector between the j -th neuron of the hidden layer and the neurons of the input layer. The weight vectors among all neurons of hidden layer and all neurons of input layer can form a weight matrix as follows:

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{bmatrix}. \quad (9)$$

$v_j = (v_{j1}, v_{j2}, \dots, v_{jn})$ is the weight vector between the j -th neuron of the hidden layer and the neurons of the output layer. The weight vectors between all the neurons of the hidden layer and all neurons of the output layer can form a weight matrix as follows:

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{m1} & v_{m2} & \cdots & v_{mn} \end{bmatrix}. \quad (10)$$

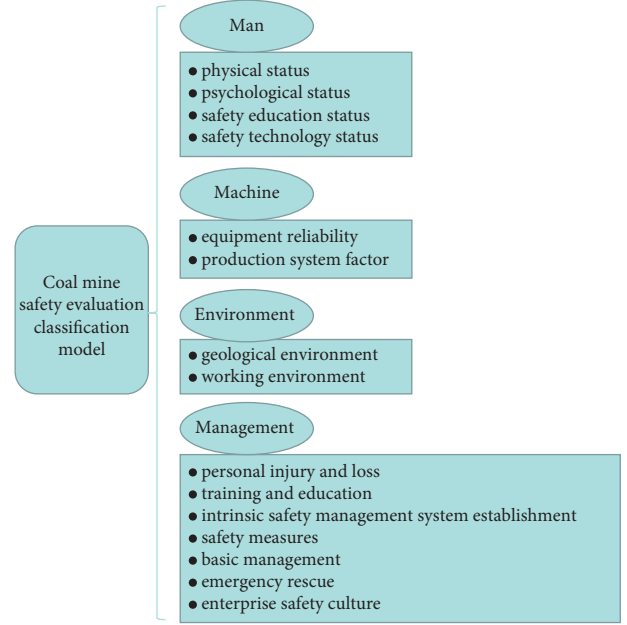


FIGURE 3: Coal mine safety evaluation classification model.

It is assumed that net_j^k represents the net input of the j -th neuron in layer k and net_j^k represents the net output of the j -th neuron in layer k . When the BP algorithm is adopted, the input-output mapping relationship of the network is given as follows:

Input Layer

$$\begin{cases} \text{net}_i^{(1)} = r_i, \\ y_i^{(1)} = \text{net}_i^{(1)}. \end{cases} \quad (11)$$

Hidden Layer

$$\text{net}_j^{(2)} = \sum_{i=1}^n w_{ji} r_i, y_j^{(2)} = f(\text{net}_j^{(2)}). \quad (12)$$

Output Layer

$$\text{net}_k^{(3)} = \sum_{j=1}^n v_{kj} y_j^{(2)}, y_k^{(3)} = \text{net}_k^{(3)}. \quad (13)$$

4.2. Application Steps. After determining the basic structure of the training sample and model, the network training and model application are carried out according to the following steps shown in Figure 4.

5. Case Study

17 evaluation indexes of 8 enterprises reflecting the state of sustainable development are selected as learning samples. All samples have been normalized according to the rules in Table 1, as shown in Table 3.

The above samples are trained through the network, and the network evaluation results are obtained, as shown in Table 4. It can be seen that the network output values of enterprise 1 and enterprise 2 are between (0.85 and 1.00),

TABLE 2: Scoring standard for coal mine safety evaluation.

Safety grade	Safety classification	Evaluation score	Characterization state
1	Intrinsic safety I	[90,100)	Safe coal mine (ideal)
2	Intrinsic safety II	[80,90)	Basic safety (good)
3	Intrinsic safety III	[70,80)	Coal mine with poor safety, early warning status (general)
4	Intrinsic safety IV	[60,70)	Unsafe coal mine, medium alarm status (poor)
5	Failure to achieve intrinsic safety	[0,60)	Unsafe coal mine, heavy alarm state (bad)

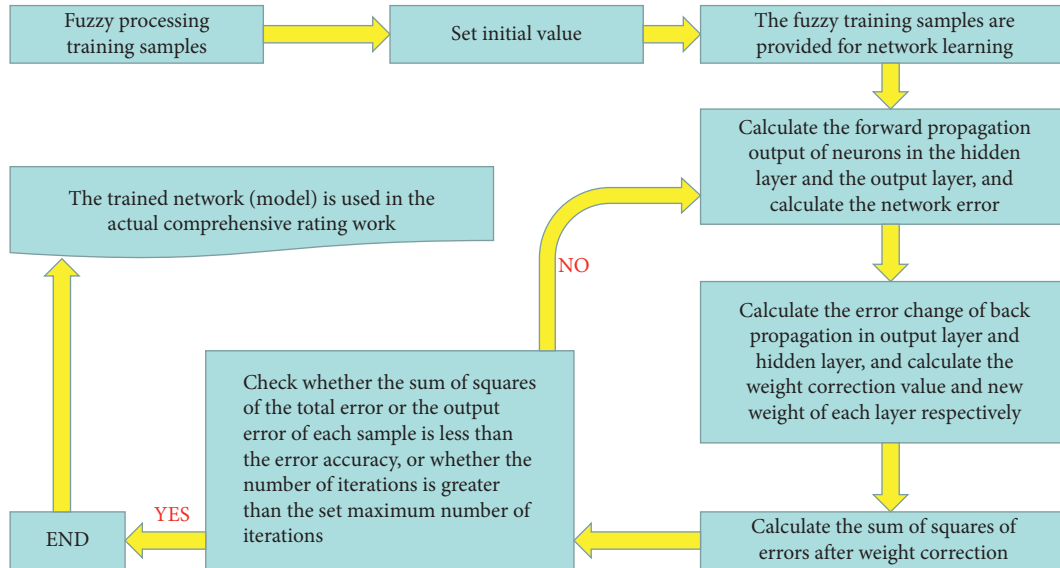


FIGURE 4: Network training and model application steps.

TABLE 3: Network learning sample.

Evaluation index	Network learning sample (enterprises)							
	1	2	3	4	5	6	7	8
Employee medical insurance rate (%)	0.85	0.80	0.74	0.71	0.63	0.63	0.52	0.50
Per capita housing area of employees (m ² /person)	0.90	0.90	0.81	0.78	0.75	0.80	0.73	0.74
Average education level of employees	0.90	0.90	0.80	0.70	0.80	0.60	0.40	0.60
Enrollment rate of employees' children (%)	0.88	0.96	0.85	0.89	0.88	0.78	0.79	0.94
Reserve production ratio (%)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Recovery ratio (%)	0.88	0.78	0.92	0.88	0.81	0.79	0.69	0.77
Comprehensive utilization degree of resources	0.80	0.90	0.80	0.60	0.70	0.80	0.80	0.60
Proportion of non-coal industry (%)	0.55	0.52	0.54	0.45	0.40	0.42	0.43	0.38
Total efficiency of raw coal (t/person)	0.63	0.62	0.56	0.70	0.55	0.52	0.44	0.45
Per capita income of miners (Yuan/month)	0.86	0.80	0.77	0.55	0.63	0.45	0.70	0.66
Ratio of profits and taxes on funds (%)	1.00	1.00	1.00	1.00	0.80	0.80	0.70	0.70
Land reclamation rate of mining area (%)	0.42	0.50	0.42	0.22	0.31	0.36	0.43	0.33
Standard rate of noise control in mining area (%)	0.88	0.85	0.77	0.80	0.72	0.68	0.65	0.50
Standard rate of wastewater treatment in mining area (%)	0.90	0.80	0.80	0.78	0.75	0.60	0.56	0.54
Standard rate of waste gas treatment in mining area (%)	0.88	0.86	0.77	0.78	0.72	0.88	0.89	0.91
Standard rate of solid waste discharge in mining area (%)	0.90	0.80	0.70	0.70	0.55	0.56	0.45	0.45
Million ton mortality (%)	1.00	1.00	1.00	1.00	1.00	1.00	0.95	0.94

enterprise 3 and enterprise 4 are between (0.70 and 0.85), enterprise 5 and enterprise 6 are between (0.50 and 0.70), and enterprise 7 and enterprise 8 are between (0.00 and 0.50). Their development levels are in sustainable development, primary sustainable development, transition from traditional development to sustainable development, and traditional development, respectively.

Based on the evaluation results of sustainable development level, the safety evaluation of enterprise 1 is now carried out. According to the intrinsic safety evaluation index system established above, data of the coal mine site are collected as shown in Table 5. Among them, the first 10 rows are the known safety assessment data of the first 10 months, the corresponding actual safety assessment value is used as

TABLE 4: Evaluation results.

Enterprises	Network output value	Sustainable development level	
1	0.8806	First level	Sustainable development
2	0.8952	First level	Sustainable development
3	0.7277	second level	Primary sustainable development
4	0.8096	second level	Primary sustainable development
5	0.5891	Third level	Transition from traditional development to sustainable development
6	0.6122	Third level	Transition from traditional development to sustainable development
7	0.3963	Fourth level	Traditional development
8	0.2293	Fourth level	Traditional development

TABLE 5: Data of coal mine safety.

	Man	Machine	Environment	Management	Expected value	Simulated value of BP neural network
1	0.89	0.91	0.81	0.85	0.90	/
2	0.88	0.91	0.86	0.88	0.90	/
3	0.82	0.86	0.84	0.81	0.91	/
4	0.85	0.82	0.80	0.79	0.88	/
5	0.84	0.89	0.81	0.80	0.87	/
6	0.88	0.85	0.82	0.86	0.81	/
7	0.85	0.82	0.83	0.86	0.89	/
8	0.85	0.86	0.88	0.82	0.87	/
9	0.86	0.85	0.92	0.86	0.91	/
10	0.90	0.93	0.85	0.91	0.93	/
11	0.86	0.93	0.88	0.91	/	0.93
12	0.85	0.95	0.81	0.93	/	0.91

the expected output as the training sample to train the BP evaluation model, and the last 2 rows are the prediction samples, representing the safety assessment data of the next 2 months.

The four elements (man, machine, environment, and management) in the coal mine and their synthesis are calculated and analyzed by using the Matlab tool and the BP neural network program. The predicted value of the personnel intrinsic safety value is (0.86–0.85). The intrinsic safety value of the equipment is (0.93, 0.95). The intrinsic safety value of environment is (0.87, 0.80). The management intrinsic safety values are (0.92, 0.93) and (0.93, 0.95). The final predicted value of coal mine intrinsic safety system is (0.93, 0.91). Therefore, it is established that the intrinsic safety degree of coal mine enterprise 1 is level I. At the same time, it also shows that the effectiveness of the neural network model applied to intrinsic safety evaluation is limited to space, and only the system intrinsic safety data is listed here as an example.

6. Conclusions

This paper establishes a three-layer BP neural network evaluation model to evaluate the sustainable development level of coal enterprises and obtains the sustainable development status of each enterprise. From the output layer neurons to the input layer neurons, the connection weights are corrected layer by layer, and the error back propagation correction is continuously implemented in the process of network training and learning, so as to reduce the error between the desired output and the actual output and improve the accuracy of the network response to the input mode. The evaluation results are completely consistent with the actual situation. The advantage of

this method is that it avoids the subjectivity and complex mathematical derivation in the traditional evaluation methods and can still get stable and correct results in the case of missing samples and parameter drift. It will provide scientific and theoretical guidance for the scientific decision-making of sustainable development of coal enterprises and has certain research value.

Based on the sustainable development evaluation of coal mines, the establishment of intrinsically safe coal mine is the development and sublimation of the existing safety management mode and coal mine safety quality standardization. It systematizes the new concept of coal mine intrinsic safety, and the established coal mine intrinsic safety evaluation system and evaluation model are applied to the coal mine site. It can provide theoretical basis and technical support for the safety management of coal mining enterprises, effectively improve the level of coal mine safety production, eliminate hidden dangers of accidents, prevent and control accidents, standardize and improve various safety management systems, and improve the safety production situation of coal mines.

Future research will focus on two aspects: (1) optimizing the processing process of the algorithm proposed in this paper to further improve the accuracy and efficiency of the algorithm; and (2) using big data technology to analyze and process the text data recorded in the process of coal mine production and comprehensively and systematically analyze the text data of coal mines to improve the risk precontrol ability of coal mine safety production.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the grant from the National Natural Science Foundation of China: Study on thermodynamic characteristics and regional heat transferring methods in the process of spontaneous combustion of coal gangue (52004211).

References

- [1] B. Yu, Y. Tai, R. Gao, Q. Yao, Z. Li, and H. Xia, "The sustainable development of coal mines by new cutting roof technology," *Royal Society Open Science*, vol. 7, no. 6, Article ID 191913, 2020.
- [2] N. Uddin, M. Blommerde, R. Taplin, and D. Laurence, "Sustainable development outcomes of coal mine methane clean development mechanism projects in China," *Renewable and Sustainable Energy Reviews*, vol. 45, pp. 1–9, 2015.
- [3] S. Hu, H. Bi, X. Li, and C. Yang, "Environmental evaluation for sustainable development of coal mining in Qijiang, Western China," *International Journal of Coal Geology*, vol. 81, no. 3, pp. 163–168, 2010.
- [4] B. C. Hasanuzzaman, "Development of a framework for sustainable improvement in performance of coal mining operations," *Clean Technologies and Environmental Policy*, vol. 21, no. 1, 2019.
- [5] H. Yan, J. Zhang, N. Zhou, and J. Chen, "The enhancement of lump coal percentage by high-pressure pulsed hydraulic fracturing for sustainable development of coal mines," *Sustainability*, vol. 11, 2019.
- [6] D. Ghosh and S. K. Maiti, "Biochar assisted eco-restoration of coal mine degraded land for UN Sustainable development goals," *Land Degradation & Development*, vol. 32, 2021.
- [7] S. S. Rai, V. M. S. R. Murthy, N. Sukesh, and A. Teja, "Operational efficiency of equipment system drives environmental and economic performance of surface coal mining—a sustainable development approach," *Sustainable Development*, vol. 29, 2020.
- [8] M. Kopacz, D. Kryzia, and K. Kryzia, "Assessment of sustainable development of hard coal mining industry in Poland with use of bootstrap sampling and copula-based Monte Carlo simulation," *Journal of Cleaner Production*, vol. 159, pp. 359–373, 2017.
- [9] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, pp. 174988–175008, 2020.
- [10] S. Zhironkin and M. Cehlár, "Coal mining sustainable development: economics and technological Outlook," *Energies*, vol. 14, no. 16, p. 5029, 2021.
- [11] H. Gao, B. An, Z. Han, and Y. Guo, "The sustainable development of aged coal mine achieved by recovering pillar-blocked coal resources," *Energies*, vol. 13, 2020.
- [12] S. Ma, C. Zhou, C. Chi, and G. Yang, "Estimating physical composition of municipal solid waste in China by applying artificial neural network method," *Environmental science & technology*, vol. 54, no. 15, pp. 9609–9617, 2020.
- [13] E. Egrioglu, U. Yolcu, E. Bas, and Z. Ali, "Median-Pi artificial neural network for forecasting," *Neural Computing & Applications*, vol. 1, pp. 1–10, 2017.
- [14] L.-h. Li, J.-c. Hang, Y. Gao, and C.-y. Mu, "Using an Integrated group decision method based on SVM, TFN-RS-AHP, and TOPSIS-CD for Cloud service supplier selection," *Mathematical Problems in Engineering*, vol. 2017, Article ID 3143502, 14 pages, 2017.
- [15] X. Chen, Y. Fang, J. Chai, and Z. Xu, "Does intuitionistic fuzzy analytic hierarchy process work better than analytic hierarchy process?" *International Journal of Fuzzy Systems*, vol. 8, pp. 1–16.
- [16] L. P. Raghav, R. S. Kumar, D. K. Raju, and R. Arvind, "Analytic hierarchy process (AHP) – swarm intelligence based flexible demand response management of grid-connected microgrid," *Applied Energy*, vol. 306, 2022.
- [17] L. Li, J. Hang, H. Sun, and L. Wang, "A conjunctive multiple-criteria decision-making approach for cloud service supplier selection of manufacturing enterprise," *Advances in Mechanical Engineering*, vol. 9, no. 3, Article ID 168781401668626, 2017.
- [18] T. Sun, X. Lv, Y. Cai, and J. Huang, "Software test quality evaluation based on fuzzy mathematics," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 1, pp. 1–11, 2020.
- [19] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets ahp, multistage weight synthesis, and promethee II," *Complexity*, vol. 2020, no. 6, pp. 1–24, 2020.
- [20] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 99, p. 1, 2020.
- [21] S. Yang, M. Guo, X. Liu, P. Wang, Q. Li, and H. Liu, "Highway performance evaluation index in semiarid climate region based on fuzzy mathematics," *Advances in Materials Science and Engineering*, vol. 2019, no. 2, pp. 1–7, 2019.
- [22] Y. Li and L. Li, "Enhancing the optimization of the selection of a product service system scheme: a digital twin-driven framework," *Strojniški vestnik - Journal of Mechanical Engineering*, vol. 66, no. 9, pp. 534–543, 2020.
- [23] X. Sun, Y. Zhou, L. Yuan, X. Li, H. Shao, and X. Lu, "Integrated decision-making model for groundwater potential evaluation in mining areas using the cusp catastrophe model and principal component analysis," *Journal of Hydrology: Regional Studies*, vol. 37, no. 3, Article ID 100891, 2021.
- [24] S. Salata and C. Grillenzoni, "A spatial evaluation of multi-functional ecosystem service networks using principal component analysis: a case of study in Turin, Italy," *Ecological Indicators*, vol. 127, Article ID 107758, 2021.
- [25] R. Chao, A. Ning, A. Jw, and D. Shang, "Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting," *Knowledge-Based Systems*, vol. 56, pp. 226–239, 2014.
- [26] W. Shouxiang, Z. Na, W. Lei, and W. Yamin, "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method," *Renewable Energy*, vol. 94, 2016.
- [27] Y. Feng and X. Xu, "A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network," *Applied Energy*, vol. 134, no. 1, pp. 102–113, 2014.
- [28] J. Xu, "Study of artificial neural network appraisal to mine coordination development ability," in *Proceedings of the International Symposiums on Information Processing*, IEEE Computer Society, St. Louis, MI, USA, April 2008.
- [29] W. Fu, Y. Xu, L. Liu, and L. Zhang, "Design and research of intelligent safety monitoring robot for coal mine shaft construction," *Advances in Civil Engineering*, vol. 2021, no. 1, 16 pages, Article ID 6897767, 2021.

- [30] K. Biao, C. Zuoyong, S. Ting, Q. Chaoxin, and Z. Yue, "Safety hazards in coal mines of Guizhou China during 2011–2020," *Safety Science*, vol. 145, Article ID 105493, 2022.
- [31] H. Wei and Y. Dong, "Adaptive fuzzy neural network control for a constrained robot using impedance learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 99, pp. 1–13, 2018.
- [32] J. Tang, F. Liu, Y. Zou, and Y. Wang, "An improved fuzzy neural network for traffic speed prediction considering periodic characteristic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, 2017.
- [33] K. Shi, J. Wang, Y. Tang, and S. Zhing, "Reliable asynchronous sampled-data filtering of T-S fuzzy uncertain delayed neural networks with stochastic switched topologies," *Fuzzy Sets and Systems*, vol. 381, 2018.

Research Article

Aided Recognition and Training of Music Features Based on the Internet of Things and Artificial Intelligence

Xidan Zhang 

Teacher College, Columbia University, New York 10027, NY, USA

Correspondence should be addressed to Xidan Zhang; xz3038@tc.columbia.edu

Received 5 January 2022; Revised 29 January 2022; Accepted 10 February 2022; Published 11 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xidan Zhang. 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 the Internet of Things, many industries have been on the train of the information age, and digital audio technology is also constantly developing. Music retrieval has gradually become a research hotspot in the music industry. Among them, the auxiliary recognition of music characteristics is also a particularly important task. Music retrieval is mainly to manually extract music signals, but now the music signal extraction technology has encountered a bottleneck. The article uses Internet and artificial intelligence technology to design an SNN music feature recognition model to identify and classify music features. The research results of the article show (1) statistic graphs of the main melody and accompanying melody of different music. The absolute value of the main melody and accompanying melody mainly fluctuates in the range of 0–7, and the proportion of the main melody can reach 36%. The accompanying melody can reach 17%. After the absolute value of the interval reaches 13, the interval ratio of the main melody and the accompanying melody tends to be stable, maintaining between 0.6 and 0.9, and the melody interval ratio value completely coincides; the main melody in the interval variable is X . (1) The relative difference value in the interval of $-X(16)$ fluctuates greatly. After the absolute value of the interval reaches 17, the interval ratio of the main melody and the accompanying melody tends to be stable, maintaining between 0.01 and 0.04 and the main melody. The value of the difference is always higher than the accompanying melody. (2) When the number of feature maps is $24 * 5$, the recognition result is the most accurate, MAP recognition result can reach 78.8, and the recognition result of precision@ is 79.2; when the feature map size is $5 * 5$, the recognition result is the most accurate, MAP recognition result can reach 78.9, the recognition result of precision@ is 79.2, and the recognition result of HAM2 (%) is 78.6. The detection accuracy of the SNN music recognition model proposed in the article is the highest. When the number of bits is 64, the detection accuracy of the SNN detection model is 59.2%, and the detection accuracy of the improved SNN music recognition model is 79.3%, which is better than the detection rate of ITQ music recognition model of 17.9%, which is 61.4% higher. The experimental data further shows that the detection efficiency of the ITQ music recognition model is the highest. (3) The SNN music recognition model proposed in the article has the highest detection accuracy, regardless of whether it is in a noisy or no-noise music environment, with an accuracy rate of 97.97% and a detection accuracy value of 0.88, which is 5 types of music. The highest one among the recognition models, the ITQ music recognition model, has the lowest detection accuracy, with a detection accuracy of 67.47% in the absence of noise and a detection accuracy of 70.23% in the presence of noise. Although there is a certain noise removal technology, it can suppress noise interference to a certain extent, but cannot accurately describe music information, and the detection accuracy rate is also low.

1. Introduction

Because the network has the advantages of fast information dissemination, easy use, and sufficient network resources, it is widely used in human work and study life. At present, with the rapid development of popular music in our country, music is everywhere, and the music wave has also affected us.

When faced with a wide variety of music types, users will inevitably feel at a loss. Users need to spend a lot of time choosing the type of music they are interested in. This method is not only a waste of time, but also very inefficient. Based on the above background, it is inevitable to design an intelligent auxiliary model of music characteristics. Literature [1] studied the ability of using self-organizing neural

mapping as a music style classifier for music fragments. The article cuts the music melody into many segments of equal length, then analyzes the music melody and rhythm, and presents the analyzed data to SOM. Document [2] discloses a system and method for implementing a simple and fast real-time single note recognition algorithm based on fuzzy pattern matching. The system can accept the music rhythm and notes during the performance, and then compared with the correct music rhythm, you can know whether the music rhythm during the performance is standard. Literature [3] proposed a new method for automatic music genre recognition in the visual domain using two texture descriptors. Literature [4] introduces the use of a dynamic set of classifier selection schemes and creates a classifier pool to perform automatic music genre classification. The working principle of the classifier is the principle of support vector machine, which can extract effective information from the spectrum image of music. The research results of the article show that the accuracy of music extraction can reach 83%. Literature [5] introduced optical music recognition technology and proposed a method for computer to automatically recognize music scores. The system can scan the printed images of music scores to extract effective information and then automatically generate audio files to provide users with listening functions. Literature [6] proposed a statistical method to deal with the task of handwritten music recognition in early notation. This method of processing music is different from the traditional method in that it directly recognizes the music signal without dividing the music signal into many paragraphs. Literature [7] investigated various aspects of automatic emotion recognition in music. Music is also a good way to express emotions. Different classifications and timbres in music will interpret different musical effects. This article explores the extensive research on music emotion recognition. Literature [8] studied the utility of the most advanced pretraining deep audio embedding method used in the task of music emotion recognition. Literature [9] proposed a music emotion recognition method based on adaptive aggregation regression model. Emotion recognition of music is an important task to evaluate the influence of music on the emotions of listeners. The article proposes an emotion estimation model, which uses the variance obtained by Gaussian process regression to measure the confidence of the estimation results of each regression model. Literature [10] proposed a new method of using template matching and pixel pattern features in computer games. The general music model does not have much to do with the change of the font, but the beats and notes of some notes do not maintain the original shape of the music signal. The model proposed in the article can be applied to these music symbols. Literature [11] proposed a method to solve the problem of multidimensional music emotion recognition, combining standard and melody audio features. Literature [12] studied the reduction of the number of training examples in music genre recognition. The article studies the impact of the reduction of training real numbers on the detection results in the process of music style recognition. The experimental results show that although the number of experiments is greatly reduced during the detection process, it can still maintain a

high classification performance in many cases. Literature [13] presents a method to parse solo performances into individual note components and use support vector machines to adjust the back-end classifier. In order to realize the generalization of instrument recognition to ready-made, commercial solo music, [14] proposed a method of musical instrument recognition in chord recording. Literature [15] proposed a method for analyzing and recognizing music speech signals based on speech feature extraction. The method is to extract effective music information from the music signal and then reorganize the music signal to a certain extent, so as to achieve the function of noise reduction. The results of the experiment show that the reorganized music signal has good noise reduction compared with the original music signal ability.

2. Research on Auxiliary Recognition of Music Features

2.1. Overall Structure of Music Feature Recognition. The music feature recognition system based on the Internet of Things technology is mainly composed of a physical perception layer, a capability layer, an adaptation layer, and a system application layer. The overall structure of the system is shown in Figure 1.

2.2. Design of Music Collection Module. To identify the music signal, it is necessary to collect the music signal first. The music collection module is composed of two parts, namely, the collection submodule and the encoding module. The music collection submodule is composed of sound sensors installed in different positions and is responsible for collecting the original music signal [16]. The sound sensor has a built-in capacitive electret microphone that is sensitive to sound, which is converted by an A/D converter and transmitted to the voice coding submodule [17]. The voice coding submodule is mainly responsible for the high-fidelity and lossless compression of the original music signal, converts the music signal into transmittable data information, and then transmits it to the music signal processing module.

2.3. Music Signal Module Processing Design. The music signal processing module is designed by a DSP processor [18]. The module uses a fixed DS chip suitable for voice signal processing. The DSP chip has low power consumption and fast running speed. It carries 2 MCBSPS, can be connected to CODEC for voice input, and has an 8-bit enhanced host parallel port to communicate with the host. Establish a communication connection, including 4 KB ROM and 16 KB DARAM. Its structure is shown in Figure 2:

3. Music Feature Assisted Recognition and Training

3.1. Extraction of Basic Music Features. Pitch, time value, and tone intensity are the most basic elements of music characteristics. The formula for the pitch level of music is defined as

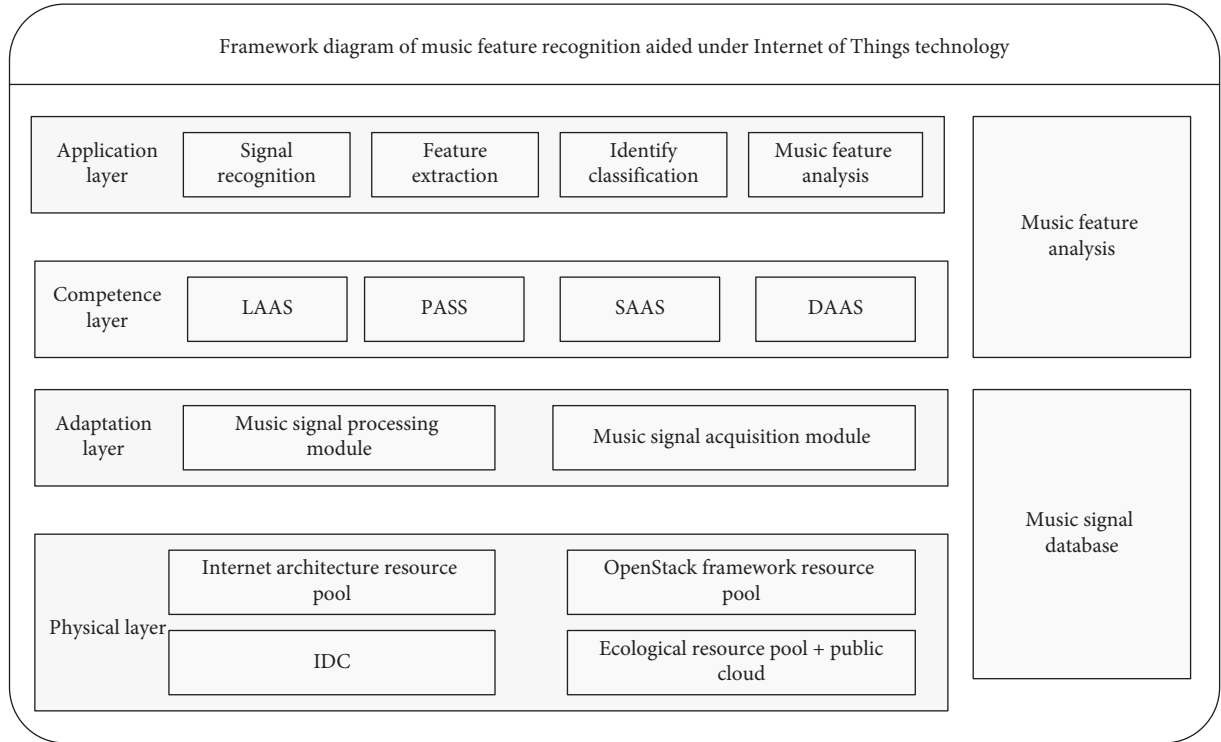


FIGURE 1: Framework diagram of music feature recognition.

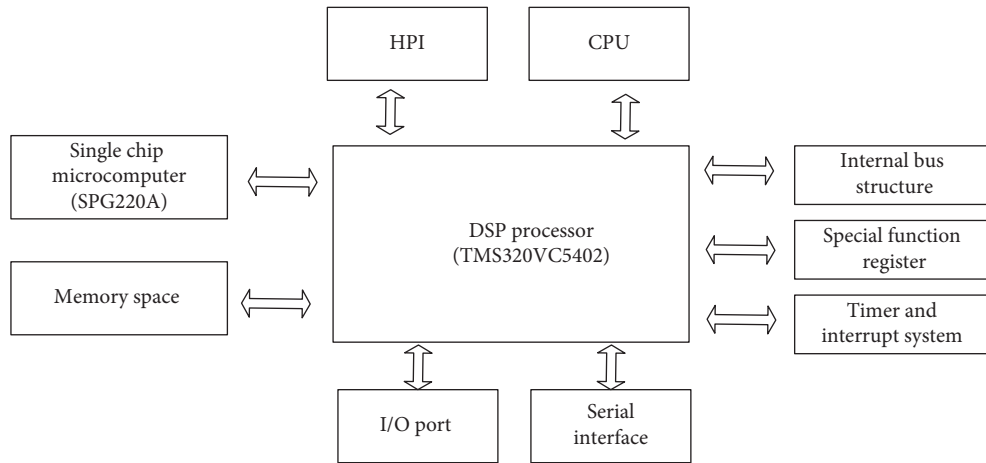


FIGURE 2: Functional structure diagram.

$$\bar{p} = \frac{\sum_{i=1}^n p_i}{n}. \quad (1)$$

p_i represents the pitch of the i note, and n represents the number of notes in the music.

Treble changes:

$$\text{Var}_-p = \frac{\sum_{i=1}^{N-1} \text{Bar}_{i+1} - \text{Bar}_i}{N\bar{P}}. \quad (2)$$

The pitch mean square error can be used to express the pitch change:

$$\text{Var}_-P = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(P_i - \frac{\sum_{i=1}^n P_i}{n} \right)^2}. \quad (3)$$

The range describes the breadth of the pitch of the music:

$$\text{range} = \text{Max}(P_1, P_2, \dots, P_n) - \text{Min}(P_1, P_2, \dots, P_n). \quad (4)$$

Time value:

$$\text{duration} = \text{end time} - \text{start time}. \quad (5)$$

3.1.1. *Tone and Music Feature Extraction.* The frequency spectrum distribution of music signals and the emotions expressed by timbre perception are shown in Table 1 [19].

The formula for extracting music strength is

$$\text{Dyn} = \frac{1}{n} \sum_{i=1}^n I_t, \quad (6)$$

$$\text{Var_Dyn} = \sqrt{\frac{1}{m} \sum_{i=1}^m \left(I_i - \frac{\sum_{i=1}^m I_i}{m} \right)^2}.$$

The degree of musical intensity change can also be expressed as

$$\text{Var_Dyn} = \frac{\sum_{i=1}^{N-1} D_{t+1} - D_t}{N \cdot \text{Dyn}}. \quad (7)$$

3.1.2. *Melody Direction Recognition.* The expression formula of music melody is

$$\text{Mel} = \sum_{i=1}^{n-1} \frac{(P_{i+1} - P_i) \cdot D_i}{D - D_n}. \quad (8)$$

D represents the total length of all notes; D_i represents the length of the i -th note [20].

The melody direction can also be expressed as

$$\text{Mel} = \sum_{i=1}^{n-1} \frac{P_{i+1} - P_i}{D_i}. \quad (9)$$

The expression formula of pronunciation point density is

$$\text{density} = \frac{n}{D}. \quad (10)$$

The change intensity of the rhythm is

$$\text{Rhy} = \sum_{i=1}^{n-1} \left| \frac{I_{i+1} - I_i}{D_i} \right|. \quad (11)$$

Music mutation degree:

$$\text{mutation} = \text{Max} \left(\frac{|\text{BarCapacity}(i) - \text{BarCapacity}(i-1)|}{\text{Max}(\text{BarCapacity}(i))} \right). \quad (12)$$

The expression of BarCapacity is [21]

$$\text{BarCapacity} = \sum_{i=0}^n \frac{K(f_0) \cdot D_i \cdot I_i}{j},$$

$$K(f_0) = \begin{cases} \frac{90}{120 - (30f_0/500)}, & \text{when } 20 \text{ Hz} < f_0 < 500 \text{ Hz} \\ 1, & \text{when } 500 \text{ Hz} < f_0 < 1000 \text{ Hz}, \\ \frac{90}{90 - (10 \times (f_0 - 1000)/4000)}, & \text{when } 1000 \text{ Hz} < f_0 < 5000 \text{ Hz}, \end{cases} \quad (13)$$

$$j = \frac{\sum_{i=0}^n \text{Velocity}_i}{120}.$$

3.2. *Musical Inference Rules.* Sudden changes in treble or tone stability appear in the sequence variance. In order to measure these change points, first express the music as the following time sequence:

$$Y_k = \mu + \varepsilon_k, \quad k = 0, 1, 2, 3, \dots, T. \quad (14)$$

Among them, μ represents the unknown constant mean value of time series Y_k , and σ^2 represents the unknown constant variance of time series Y_k (and ε_k).

Get the iterative residual sequence:

$$a_k = \frac{Y_k - \left(\sum_{i=0}^{k-1} (Y_i/k) \right)}{\sqrt{(k+1/k)S_y^2}}, \quad k = 0, 1, 2, 3, \dots, T, \quad (15)$$

and make

$$C_k = \sum_{i=1}^k a_i^2, \quad k = 0, 1, 2, 3, \dots, T. \quad (16)$$

Get statistics:

$$W_k = \frac{C_k}{C_T}, \quad k = 0, 1, 2, 3, \dots, T. \quad (17)$$

After centralized processing,

$$D_k = \frac{C_k}{C_T} - \frac{k}{T}, \quad k = 0, 1, 2, 3, \dots, T, D_0 = D_T = 0. \quad (18)$$

3.3. *Music Separation Algorithm.* According to the difference between the impact sound and the harmonic sound in the frequency spectrum, we can separate the original spectrum

TABLE 1: Tone-emotion mapping relationship.

Musical emotion type	Tonal characteristics
Hate class	The tone is sharp, rough, and bright
Depression	The tone is deep, pure, simple, monotonous bass, dim, plain, and hollow
Calm meditation class	The tone is soft, pure, and simple
Desire	The tone is deep, plain, hollow, pure, and simple
Pastoral style	The tone is soft, pure, and simple
Perceptual	The tone is soft, sweet and soft, rich, gorgeous, pleasant, and nasal
Active class	The tone is bright, rich, and gorgeous
Awesome	The tone is bright, sharp, and brilliant

W_{fJ} into the impact spectrum P_{fJ} and the harmonic spectrum, which is

$$W_{fJ} = P_{fJ} + H_{fJ}. \quad (19)$$

The separation of impact sound and harmonic sound:

$$Q(H^t, P^t, U^t, V^t) = \frac{1}{\sigma_H^2} \sum_{fJ} \left\{ (H_{f,t-1}^t - U_{f,t}^t)^2 - (H_{f,t}^t - U_{f,t}^t)^2 \right\} \\ + \frac{1}{\sigma_P^2} \sum_{fJ} \left\{ (P_{f,t-1}^t - V_{f,t}^t)^2 - (P_{f,t}^t - V_{f,t}^t)^2 \right\}. \quad (20)$$

Minimum:

$$H_{f,J}^{t+1} = H_{fJ}^t + \Delta^t, \quad (21) \\ P_{f,J}^{t+1} = P_{fJ}^t + \Delta^t,$$

in

$$\Delta^t = \frac{\alpha}{4} (H_{fJ-1}^t - 2H_{f,t+1}^t + H_{f,t+1}^t) - \frac{1-\alpha}{4} \\ (P_{fJ-1}^t - 2P_{f,t+1}^t + P_{f,t+1}^t), \quad (22) \\ \alpha = \frac{\sigma_y^2}{\sigma_H^2 + \sigma_y^2}.$$

4. Simulation Experiment

4.1. Music Feature Recognition

4.1.1. *Algorithm Definition.* Algorithm definition is as shown in Table 2.

4.1.2. *Experimental Data and Research.* The article uses the Internet of Things and human intelligence technology to design an SNN music feature assisted recognition model. In order to detect the recognition efficiency of the SNN music feature assisted recognition model, the experiment selected more than 50 pieces of multiple types of music for music feature recognition and counted them separately: the main melody and accompanying melody curves of different music. The main melody lines of different types of music are different, and the main characteristics of music melody are linearity and fluidity. The abscissa of the experimental

statistics graph represents the absolute value of the interval, and the ordinate represents the percentage of the absolute value of the interval. The specific experimental results are shown in Figure 3.

From the data in Figure 3, we can conclude that the absolute value of the interval between the main melody and the accompanying melody mainly fluctuates in the range of 0–7. In the interval line chart of the main melody, the second degree accounted for the highest proportion of the interval melody. Reaching 36%, in the interval line chart accompanying the melody, 5 degrees accounted for the highest proportion of interval melody, up to 17%. After the absolute value of the interval reaches 13, the interval ratio of the main melody and the accompanying melody tends to be stable, maintaining between 0.6 and 0.9, and the melody interval ratio values completely coincide.

According to the experimental data in Figure 4, we can conclude that the relative difference of the main melody within the interval of $X(1)$ – $X(16)$ fluctuates greatly. When the interval variable is $X(3)$, the relative difference is the largest. The maximum can reach 0.79. The relative difference value of the accompanying melody in the interval variable $X(1)$ – $X(10)$ fluctuates greatly. When the interval variable is $X(3)$, the relative difference value is the largest, and the maximum can reach 0.61. After the absolute value of the interval reaches 17, the interval ratio of the main melody and the accompanying melody tends to be stable, maintaining between 0.01 and 0.04, and the difference between the main melody and the accompanying melody is always higher than that of the accompanying melody.

4.2. Comparative Experiment and Analysis

4.2.1. *The Influence Experiment of Feature Map.* Based on the same recognition results of different features, it can directly reflect the recognition accuracy of different models and experimentally study the influence of the number and size of feature maps on the detection results. In the different feature map number recognition experiment, different distributions of convolutional layers were selected, and the distribution size was from 8 to 64. In the different feature map size recognition results experiment, 11 feature maps of different sizes were selected. The experimental data is shown in Tables 3 and 4.

According to the data in Table 3 and Figure 5, we can conclude that when the number of feature maps is $24 * 5$, the recognition result is the most accurate, the MAP recognition

TABLE 2: Algorithm definition table.

	Definition	Publicity
Interval statistics	Statistics of the melody interval are carried out on each track of the music [22]	$Interval_Sta_i = \begin{cases} Interval_i , & Interval_i < 25 \\ 25, & Interval_i \geq 25 \end{cases}$
Classification algorithm	According to the different characteristics of the interval distribution, the main track and the accompanying track are distinguished [23]	$Rhythm = (n/T) = (n/(Dura_Meter \times Dura_Num))$

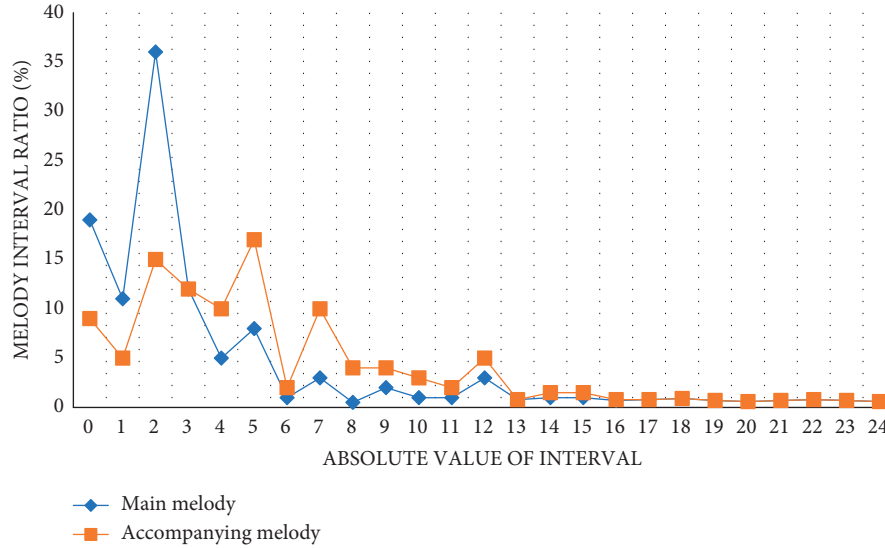


FIGURE 3: Interval statistics of different audio tracks.

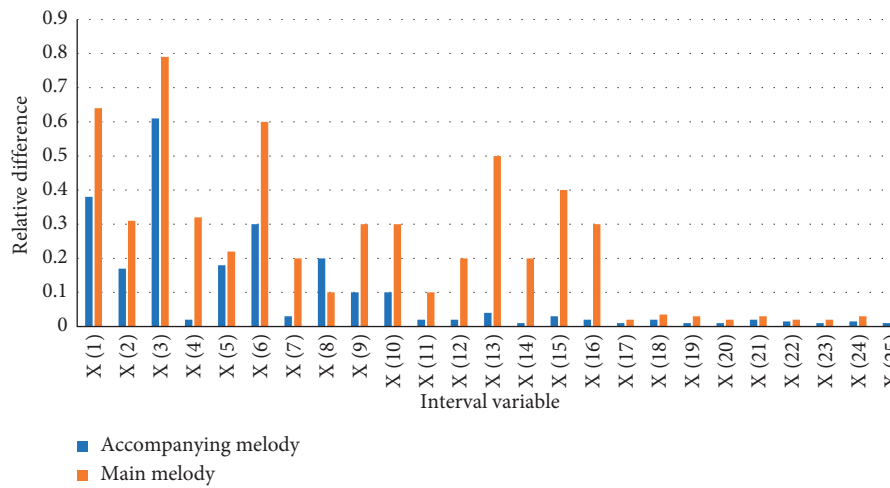


FIGURE 4: Relative difference distribution of interval statistics.

TABLE 3: Recognition results of different number of feature maps.

Quantity	MAP (%)	Precision@500 (%)	HAM2 (%)
8 * 5	74.7	76.3	74.9
16 * 5	74.5	77.1	76.2
24 * 5	78.8	79.2	79.6
32 * 5	77.6	78.9	78.1
48 * 5	77.3	77.5	76.6
64 * 5	74.8	76.6	75.7

TABLE 4: Recognition results of different feature map sizes.

Quantity	MAP (%)	Precision@500 (%)	HAM2 (%)
4 * 4	78.7	79.2	78.8
5 * 5	78.9	79.2	78.6
6 * 6	78.8	79.2	79.6
7 * 7	78.7	78.9	78.6
8 * 8	77.6	77.9	77.7
9 * 9	76.4	77.2	76.8
10 * 10	76.8	77.5	77.1
11 * 11	75.7	76.6	76.3
12 * 12	75.8	75.9	76.1
13 * 13	75.2	75.6	75.7
14 * 14	74.1	75.7	75.8

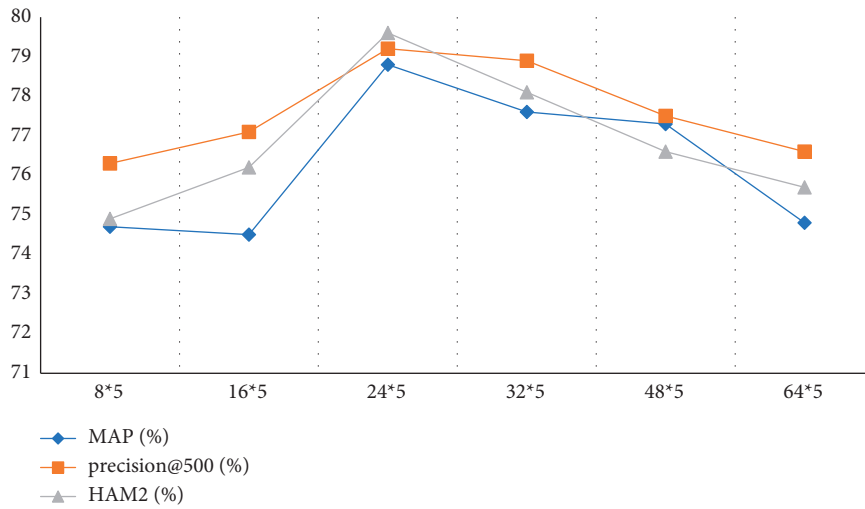


FIGURE 5: Statistics of recognition results.

result can reach 78.8, the recognition result of precision@ is 79.2, and the recognition result of HAM2 (%) is 79.6. When the number of feature maps is 8 * 5, the accuracy of the recognition result is the lowest. The recognition rate of MAP is 74.7, the recognition result of precision@ is 76.3, and the recognition result of HAM2 (%) is 74.9. In general, the detection accuracy of 6 different numbers of feature maps generally maintains above 74%.

According to the data in Table 4 and Figure 6, we can conclude that when the feature map size is 5 * 5, the recognition result is the most accurate. The MAP recognition result can reach 78.9, the recognition result of precision@ is 79.2, and the recognition result of HAM2 (%) is 78.5. When the feature map size is 78.6 and the feature map size is 14 * 14, the recognition accuracy is the lowest. The recognition accuracy of MAP is 74.1, the recognition result of precision@ is 75.7, and the recognition result of HAM2 (%) is 75.8. In general, the detection accuracy of 11 different sizes of feature maps generally maintains above 74%.

4.2.2. Comparison with Other Methods. In order to test the performance of the music recognition model, the experiment improved the SNN music recognition model proposed in the article and compared it with the detection

performance of the other three models. The experiment chose 5 different types of bit numbers. The number of bits is a unit, and the same as the sampling accuracy, the higher the baud rate or bit rate is, the more detailed the light changes of the music can be reflected. Observe the detection accuracy rates of 5 different models under different types of bits. The specific experimental data are shown in Table 5.

According to the data in Table 5 and Figure 7, we can conclude that the detection accuracy of the SNN music recognition model proposed in the article is the highest among 5 different music recognition models. When the number of bits is 64, the SNN detection accuracy rate of the improved SNN music recognition model is 59.2%, and the detection accuracy rate of the improved SNN music recognition model is 79.3%, which is 61.4% higher than the 17.9% detection rate of the ITQ music recognition model. The experimental data further shows that the ITQ music recognition model has the highest detection efficiency, which greatly promotes the efficiency of music feature auxiliary recognition.

4.3. Test Model Performance Comparison Test

4.3.1. Evaluation Criteria. The evaluation criteria are as shown in Table 6.

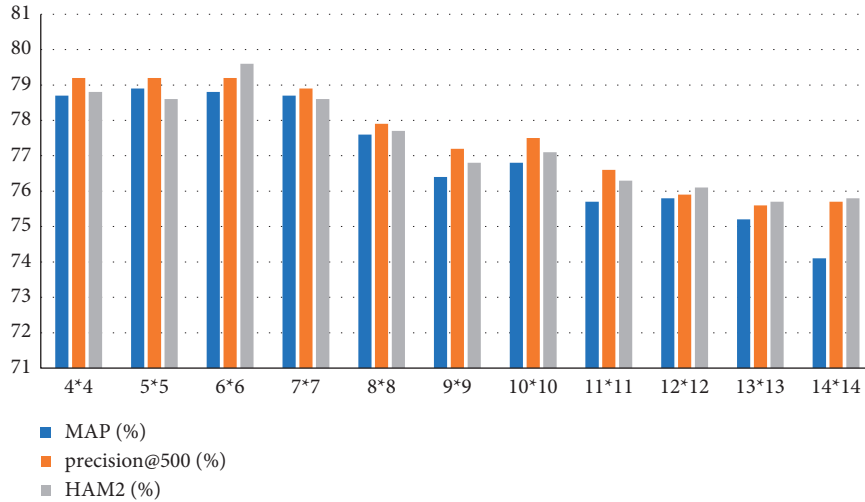


FIGURE 6: Statistics of recognition results.

TABLE 5: Average mean precision of different number of bits.

Method	16 bits	24 bits	32 bits	48 bits	64 bits
SNN music recognition model	55.2	56.6	55.8	58.1	59.2
Improved SNN music recognition model	76.7	77.9	78.3	78.9	79.3
CNNH music recognition model	46.5	52.1	52.1	53.2	53.3
KSH music recognition model	30.3	33.7	34.7	35.6	36.5
ITQ music recognition model	16.2	16.9	17.3	17.5	17.9

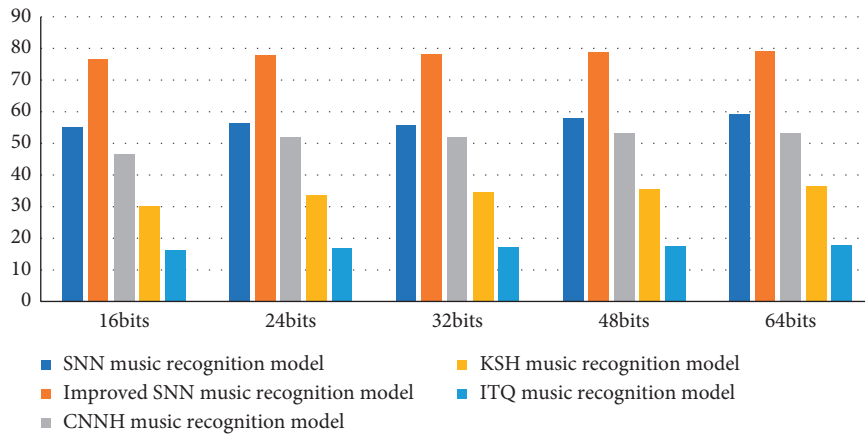


FIGURE 7: Average mean precision statistics.

TABLE 6: Evaluation criteria table.

Index	Metrics	Formula
Accuracy	The accuracy measurement standard refers to the ratio of the number of correct music types to the number of all music types [24]. The larger the index value, the more accurate the recognition result	$\text{Precision} = (\text{hits}_u / \text{recset}_u)$
Recall rate	The recall rate standard refers to the proportion of the theoretically largest number of hits that recognize musical characteristics [25]. The larger the index value, the more accurate the recognition result	$\text{Recall} = (\text{hits}_u / \text{testset}_u)$
$F1$ measurement	The $F1$ measurement index can effectively balance the accuracy rate and the recall rate by favoring the smaller value. The larger the index value, the more accurate the recognition result	$F1 = (2 \times \text{precision} \times \text{recall} / (\text{precision} + \text{recall}))$

TABLE 7: Music classification and detection object table.

Music type number	Noisy	No noise
1	10	30
2	10	30
3	20	40
4	10	40
5	20	50
6	20	50

TABLE 8: No-noise recognition results.

Model	Accuracy (%)	Accuracy (%)	Recall rate (%)	F1 score (%)
SNN music recognition model	95.71	96.83	96.74	95.82
Improved SNN music recognition model	97.97	98.10	97.89	98.00
CNNH music recognition model	80.21	82.31	83.24	84.51
KSH music recognition model	72.14	73.46	90.26	91.32
ITQ music recognition model	67.47	68.24	66.76	67.12

TABLE 9: Noisy recognition results.

Model	Accuracy (%)	Accuracy (%)	Recall rate (%)	F1 score (%)
SNN music recognition model	92.12	91.83	91.64	91.28
Improved SNN music recognition model	93.91	94.21	94.74	94.62
CNNH music recognition model	75.32	77.23	76.34	77.21
KSH music recognition model	68.62	69.24	69.12	68.24
ITQ music recognition model	70.23	71.22	74.21	72.45

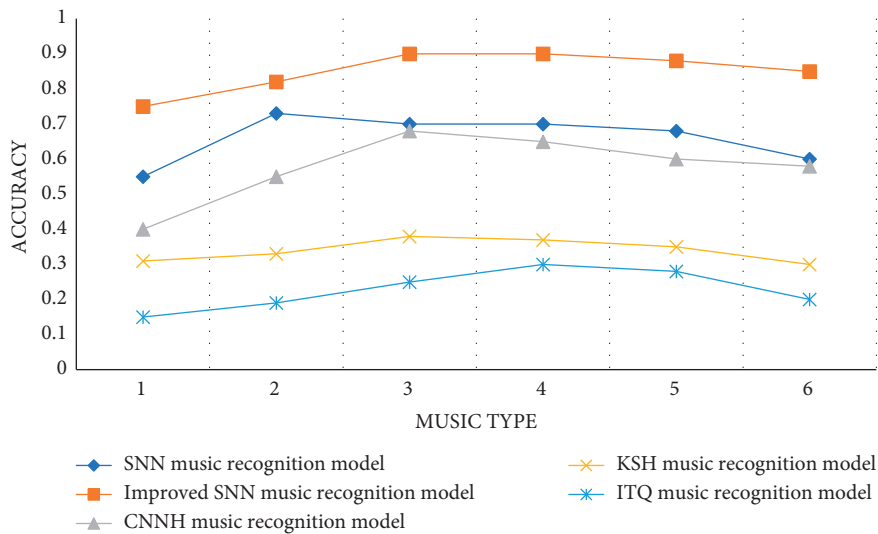


FIGURE 8: Comparison of uninteresting music classification and detection accuracy.

4.3.2. Experimental Results and Analysis. In order to test the performance of the SNN music feature-assisted recognition model, we run the model proposed in the article and other music recognition models under noisy and no-noise music conditions, observe the detection accuracy of different models, and verify the detection accuracy of different models. In order to make the experimental results more analytical, we have selected 5 different types of music data. The experiment detects these 5 types of music data with and

without noise and observes the experimental results. The music sample data is shown in Table 7, and the specific detection results are shown in Tables 8 and 9.

According to the data in Table 8 and Figure 8, we can conclude that the SNN music recognition model proposed in the article has the highest detection accuracy, with an accuracy rate of 97.97% and a detection accuracy value of 0.88. It is 5 types of music recognition models. The tallest one among them, the ITQ music recognition model, has the

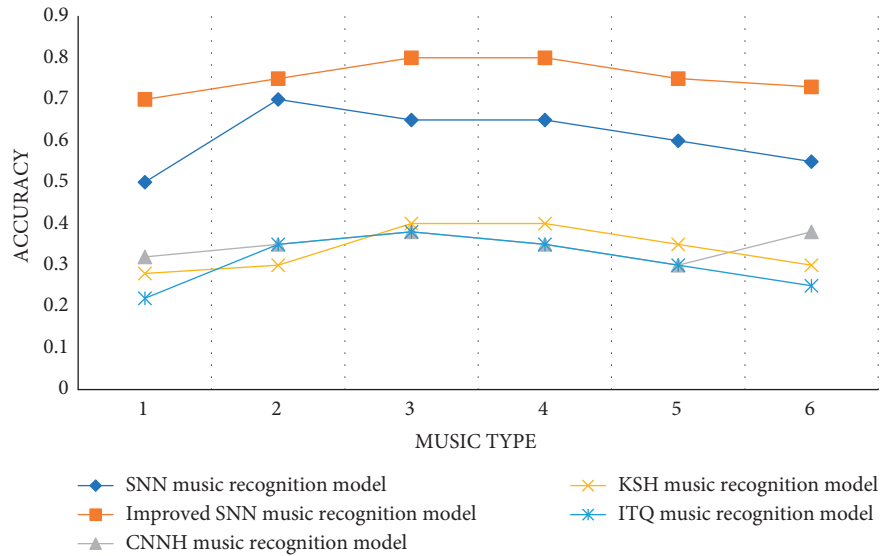


FIGURE 9: Comparison of dry music classification and detection accuracy.

lowest detection accuracy rate of 67.47%, and the highest detection accuracy value is 0.3. The CNNH music recognition model and the KSH music recognition model are in the middle of the highest value and the lowest value.

We can find from Figure 9 that the ITQ music recognition model has the lowest detection accuracy. The detection accuracy rate in the absence of noise is 67.47%, and the detection accuracy in the presence of noise is 70.23%. Although there is a certain noise removal technology, it can suppress noise interference to a certain extent, but cannot accurately describe music information, and the detection accuracy rate is also low. The detection accuracy of the KSH music recognition model is higher than that of the ITQ music recognition model, which can accurately describe the changes of music signals, but there are certain defects in noise processing, and the error rate of music detection is relatively large. The SNN music feature assisted recognition model proposed in the article has the highest detection accuracy among the five models, and there are many types of music detected, which can analyze music signals more comprehensively and systematically, and the accuracy rate is as high as 99.12%, thus greatly improving the efficiency of music detection. It is believed that the detection accuracy will be improved by using feature extraction approaches.

5. Conclusion

Today we are in an era of informationization and intelligence. The use of intelligent methods to study music has attracted more and more people's attention. Computer music has also made many achievements and has a very broad market prospect. Using a computer to simulate music signals, this process involves not only computers and music, but also a lot of complex professional knowledge. At present, there are still many problems in the auxiliary recognition of music characteristics in human intelligence; despite the design of the auxiliary model of music characteristics recognition in the article, music signals can be analyzed and

identified efficiently, but the way of expressing music needs further research.

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 regarding this work.

References

- [1] P. J. P. D. León and J. Quereda, "Feature-driven recognition of music styles," in *Proceedings of the Pattern Recognition and Image Analysis, First Iberian Conference, Ib PRIA 2003*, pp. 51–65, DBLP, Puerto de Andratx, Mallorca, Spain, June 2003.
- [2] M. S. Sinith, S. Tripathi, and K. V. V. Murthy, "Real-time swara recognition system in Indian Music using TMS320C6713," in *Proceedings of the 2015 International Conference on Advances in Computing*, pp. 51–65, Communications and Informatics (ICACCI), Kochi, India, 2015.
- [3] Y. Costa, L. Oliveira, and A. Koerich, "Music genre recognition using gabor filters and LPQ texture descriptors," in *Proceedings of the Iberoamerican Congress on Pattern Recognition*, pp. 114–121, Springer, Berlin Heidelberg, 2013.
- [4] Y. Costa, L. Oliveira, and A. Koerich, "Music genre recognition based on visual features with dynamic ensemble of classifiers selection[J]," *IEEE*, vol. 12, no. 08, pp. 21–32, 2013.
- [5] X. Liu, "Application and research on optical music recognition," *Computer Engineering*, vol. 12, no. 11, pp. 21–36, 2003.
- [6] J. Calvo-Zaragoza, A. H. Toselli, and E. Vidal, "Early handwritten music recognition with hidden markov models[C]," *IEEE in Proceedings of the International Conference on Frontiers in Handwriting Recognition*, vol. 04, no. 08, pp. 21–31, 2017.

- [7] S. Janani, K. Iyswarya, and L. A. Visuwasam, "Critical survey on music emotion recognition techniques for music information retrieval[J]," *Progress in Textile Science & Technology*, vol. 26, no. 1, pp. 11–17, 2011.
- [8] E. Koh and S. Dubnov, "Comparison and analysis of deep audio embeddings for music emotion recognition," *Acoustic Engineering*, vol. 04, no. 12, pp. 117–121, 2021.
- [9] S. Fukayama and M. Goto, "Adaptive aggregation of regression models for music emotion recognition," *Journal of the Acoustical Society of America*, vol. 140, no. 4, p. 3091, 2016.
- [10] L. Ki Woong and C. Bong, "The music score recognition system of the robust music symbols distortion for computer games," *Journal of The Korean Society for Computer Game*, vol. 28, no. 4, pp. 17–26, 2015.
- [11] B. Rocha, R. Panda, and P. P. Rui, "Dimensional music emotion recognition: combining standard and melodic audio features," in *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research – CMMR'2013*, pp. 21–31, 2014.
- [12] I. Vatulkin, M. Preuß, and G. Rudolph, *Training Set Reduction Based on 2-Gram Feature Statistics for Music Genre Recognition*, pp. 45–52, Technische Universität, Faculty of Computer Science, Algorithm Engineering, 2008.
- [13] K. Patil and M. Elhilali, "Biomimetic spectro-temporal features for music instrument recognition in isolated notes and solo phrases[J]," *EURASIP Journal on Audio Speech and Music Processing*, vol. 2015, no. 1, pp. 1–13, 2015.
- [14] I. Vatulkin, A. Nagathil, and W. Theimer, *Performance of Specific vs. Generic Feature Sets in Polyphonic Music Instrument Recognition*, pp. 14–16, Springer, Berlin, Heidelberg, 2013.
- [15] Y. U. Li Juan, "Study on the music recognition method based on voiceprint recognition[J]," *Automation & Instrumentation*, vol. 04, no. 16, pp. 25–32, 2018.
- [16] C. Guobin, Z. Sun, and L. Zhang, "Road Identification Algorithm for Remote Sensing Images Based on Wavelet Transform and Recursive Operator," *IEEE Access*, vol. 8, pp. 141824–141837, 2020.
- [17] X. Ning, W. Li, B. Tang, and H. He, "BULDP: biomimetic uncorrelated locality discriminant projection for feature extraction in face recognition," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2575–2586, 2018.
- [18] L. L. Y. Hong and Z. Hong-Jiang, "A new approach to query by humming in music retrieval," in *Proceedings of the 2001 IEEE International Conference on Multimedia and Expo*, pp. 322–324, Tokyo, Japan, August 22–25, 2001.
- [19] D. Liu, N. Zhang, and H. Zhu, "Form and mood recognition of Johann Strauss's waltz centos," *Chinese Journal of Electronics*, vol. 12, no. 4, pp. 587–593, 2003.
- [20] D. Liu, N. Zhang, and H. Zhu, "CAD system of music animation based on form and mood recognition," *Pattern Recognition and Artificial Intelligence*, vol. 16, no. 3, pp. 271–283, 2003.
- [21] N. Juslin and Laukka, "Improving emotional communication in music performance through cognitive feedback," *Musicae Scientiae*, vol. 12, no. 08, pp. 151–183, 2004.
- [22] N. Juslin and G. Madison, "The role of timing patterns in recognition of emotional expression from musical performance Music perception," *Music Perception: An Interdisciplinary*, vol. 17, no. 2, pp. 197–221, 1999.
- [23] G. Chen, Y. Zhang, and S. Wang, "Hyperspectral remote sensing IQA via learning multiple kernels from mid-level features[J]," *Signal Processing: Image Communication*, vol. 83, p. 115804, 2020.
- [24] D. Yu, S. Wang, and L. Deng, "Sequential labeling using deep-structured conditional random fields[J]," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 6, pp. 965–973, 2010.
- [25] C. Wu, "Application of digital image based on machine learning in media art design," *Computational Intelligence and Neuroscience*, vol. 2021, p. 8546987, 2021.

Research Article

Collaborative Filtering Recommendation of Music MOOC Resources Based on Spark Architecture

Lifu Wang 

Music Conservatory, Shandong University of Arts, Jinan 250000, China

Correspondence should be addressed to Lifu Wang; z00310@sdca.edu.cn

Received 8 December 2021; Revised 15 January 2022; Accepted 9 February 2022; Published 7 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Lifu Wang. 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 MOOC platforms, MOOC resources have grown substantially, causing the problem of information overload. It is difficult for users to select the courses they need from a large number of MOOC resources. It is necessary to help users select the right music courses and at the same time make the outstanding music courses stand out. Recommendation systems are considered a more efficient way to solve the information overload problem. To improve the accuracy of the recommendation results of music MOOC resources, a mixed collaborative filtering recommendation algorithm based on Spark architecture is proposed. First, the user data and item data are modeled and scored by the collaborative filtering algorithm, then the tree structure of the XGBoost model and the features of regular learning are combined to predict the scores, and then the two algorithms are mixed to solve the optimal objective function to obtain the set of candidate recommendation data. Then, the frog-jumping algorithm is used to train the weighting factors, and the optimal combination of weighting factors is used as the training result of the samples to realize the data analysis of the mixed collaborative filtering recommendation algorithm. The experimental results in the music MOOC resource show that the average absolute error and root mean square error of the proposed method are 0.406 and 1.117, respectively, when the sparsity is 30%, which are lower than those of other existing collaborative filtering recommendation methods, with higher accuracy and execution efficiency.

1. Introduction

In recent years, with the rapid development and popularization of the World Wide Web and smart devices, all kinds of information emerge one after another. Faced with surging information, it is difficult for people to efficiently and accurately obtain the information they want from the vast sea of Internet information, which is the problem of information overload [1–5]. Internet giants such as Alibaba, Baidu, and Google all process and generate petabytes of data every day, and the Internet has also moved from the IT (Information Technology) era to the DT (Data Technology) era [6–8].

The recommendation system of Massive Open Online Course (MOOC) based on big data platform has attracted more and more attention from academia and industry. Thanks to the advantages of easy use, low cost (mostly free of charge), wide coverage of people, autonomous learning, and abundant learning resources, MOOC has developed rapidly

since its appearance [9–12]. However, with the rapid development of the MOOC platform, the number of courses has greatly increased, and the quality is uneven, which is difficult to guarantee. For users, it is difficult to accurately and efficiently find the part of courses they really like from a large number of courses, but it reduces users' experience and satisfaction with the MOOC platform [13–15]. It is difficult for the course teacher to make the course stand out and reduces the enthusiasm of the course teacher. In the long run, the user growth rate of the MOOC platform will stagnate, and even users will be lost. Quality courses will be increasingly scarce, which will affect the development of the MOOC platform. Therefore, with the development of the MOOC platform, MOOC platform also has the problem of curriculum information overload, especially music MOOC.

In order to solve the problem of information overload, scientists have put forward many solutions. Recommendation system is proved to be an effective tool to solve the

problem of information overload. According to the user's past behavior records, it uses algorithms to recommend new items that may be of interest to users to solve the problem of information overload. Commonly used recommendation algorithms mainly include content-based (CB) recommendation algorithm [16], collaborative filtering (CF) algorithm [17], latent factor model (LFM) [18], and hybrid recommendation (HR) algorithm [19]. At present, there are many research studies on personalized recommendation service for dynamic users by collaborative filtering algorithm. For example, Lim et al. [20] applied the CF algorithm to the tourism recommendation system, providing effective tourist routes and scenic spots according to users' preferences and using the user image tag as an assistant for effective recommendation. Wang et al. [21] used the XGBoost algorithm to predict and analyze the churn of Internet customers. These two algorithms are used in both prediction and recommendation systems, but there are few related research studies on combining the two algorithms.

The architecture of the traditional personalized recommendation system cannot effectively store and analyze large datasets, so it cannot guarantee the timeliness and accuracy of the recommendation model. Therefore, it is very important to study a personalized recommendation system that can store and process large datasets. At present, there are many personalized recommendation systems for big data processing, which are basically distributed frameworks. Based on the distributed framework, Soni et al. [22] and others put forward the collaborative filtering algorithm of items based on Hadoop. When recommendation technology is combined with big data technology, the performance of the big data framework will greatly affect the overall recommendation performance. In terms of sparsity and real time, Hadoop, the current mainstream big data platform, is not suitable for low-latency accurate online computing due to its design characteristics. Spark is a new generation of computing framework that emerged in recent two years. The memory-based features make it much better than the MapReduce framework in computing efficiency. From the storage point of view, the HDFS framework in Hadoop ecological environment is mainly used at present. Babu et al. [22] proposed an electronic product recommendation system based on Spark and achieved good running effect. Sundari et al. [23] proposed an integrated sentiment analysis method of mixed collaborative filtering method in a big data environment.

Therefore, this paper attempts to combine the CF algorithm with the Spark platform with faster computing speed to solve the problem of accurate personalized recommendation of music MOOC resources. The research goal is to improve the traditional CF algorithm according to the characteristics of the MOOC platform and use it as the recommendation engine of the Spark system to design a music MOOC recommendation system architecture so that the MOOC platform can provide more accurate music course recommendations to more users in the case of big data. The main work has been completed in two aspects: (1) according to the application scenario characteristics of the music MOOC platform, the distributed Spark

recommendation platform architecture is constructed and (2) combining the efficiency of the XGBoost algorithm with the effectiveness of the collaborative filtering algorithm, the dynamic personalized recommendation of users on the big data platform is completed, and the weight factors are optimized by the shuffled frog leading algorithm.

2. Distributed Spark Recommendation Platform Architecture

Recommendation systems are generally designed based on specific usage scenarios, such as Amazon Mall push system based on e-commerce, Netflix based on movie and video recommendation, Pandora based on music network radio, and Facebook based on social network. Because the data processed by the bottom layer of the recommendation system is often massive, in order to run stably in the big data environment, the latest big data computing framework Spark is adopted. In Spark architecture, data, machines, and services can be linearly increased, thus providing real-time and reliable recommendations to users [24].

2.1. Ideas of Spark Architecture. In order to face massive data, recommendation engines are often designed according to offline and online computing, so as to achieve faster, more reasonable, and more personalized recommendation. The schematic diagram of the calculation method of Spark architecture is shown in Figure 1.

The data of music MOOC resources are huge, and it is characterized by "write once and read many times" in business. In this case, the traditional relational database is not suitable. When designing a Spark-based recommendation system, this paper adopts the scheme of combining HDFS with master-slave node and realizes a high-performance data warehouse to meet the query and analysis requirements of the recommendation system in a big data environment. Specifically, the data required by the recommendation system are stored in HDFS in the form of a master-slave node, and Spark SQL is used to query the data. For master-slave node mode, Spark SQL can use SQL statements to query just like querying database table. In addition, because Spark SQL is based on Spark, it is much faster than Hive. Only the corresponding columns need to be scanned when querying, so this data warehouse design has high performance. The data warehouse design of the master-slave node is shown in Figure 2.

The data in the data warehouse contain three layers: the raw data layer, the offline intermediate layer, and the recommendation result layer. Each layer has its corresponding computation, and the amount of data and computation decreases gradually from the raw data layer to the recommendation result layer.

2.2. Overall Platform Architecture. This system adopts Scala language and is implemented based on Spark. Spark is a distributed computing framework for big data, and its running framework consists of Driver and Executor. When we submit tasks to the Spark cluster, the machine will start a

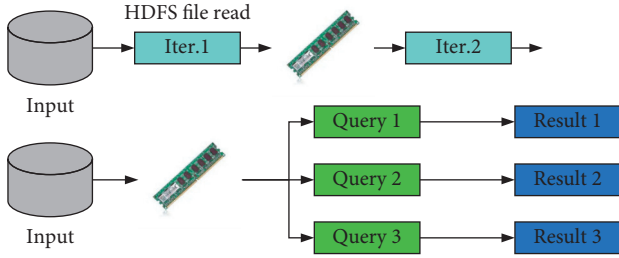


FIGURE 1: The schematic diagram of the calculation method of Spark architecture.

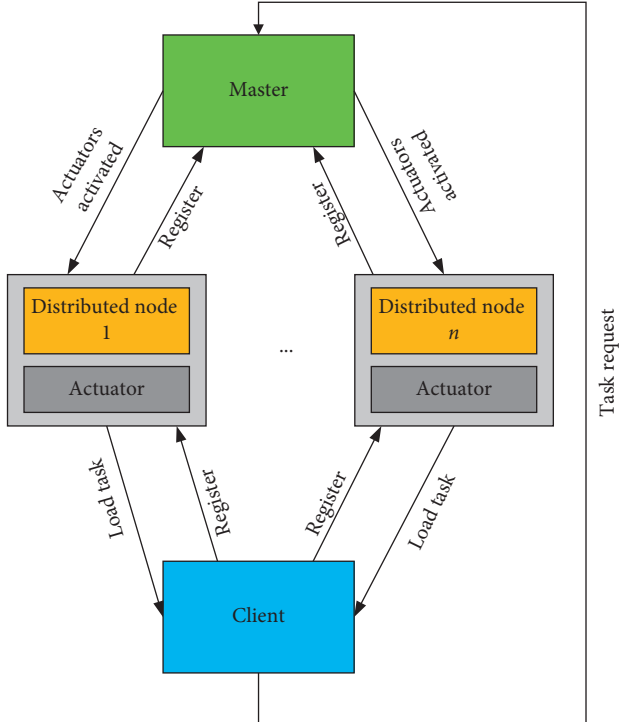


FIGURE 2: Data warehouse design with master-slave nodes.

JVM virtual machine to run Spark Context, which is called Driver. When the Spark program runs map, reduce, filter, and other operators, the functions in the operators will be submitted to each machine in the cluster for distributed operation, and the process started by each operation node is called the Executor. The distributed big data computing framework is shown in Figure 3.

Two tasks, offline and online computing, are usually run in Spark Executor, in which online computing often caches RDD data of intermediate results in memory to speed up distributed computing. The data calculated by the Executor will be returned to Spark Driver for further online processing. Spark platform architecture is shown in Figure 4.

3. Music MOOC Resource Recommendation Based on Mixed Collaborative Filtering

3.1. Mixed Collaborative Filtering. In order to improve the effectiveness and accuracy of collaborative filtering results

and to strengthen the applicability of the personalized recommendation system, the score of the traditional CF algorithm and the score of the XGBoost recommendation algorithm are weighted. According to their common scoring results, it is regarded as the limited order of personalized recommendation for dynamic users in big data. XGBoost algorithm is a learning system based on tree structure [25]. Compared with commonly used advanced algorithms, such as ant colony algorithm and fish swarm algorithm, the XGBoost algorithm has good scalability and scalability. In the distributed computing of big data, faced with the correspondence of tens of millions of users' data, XGBoost has solved the problems of memory limitation and slow algorithm speed in the time-consuming link of similarity calculation. The following is a scoring model for user data. The specific methods are as follows.

Firstly, a list of users and items is established, and each user u corresponds to a database record, which stores the situation of user u accessing items in the big data platform. In this record, the number of visits of user u to some items i can be mined. Understand user preferences according to the number of visits. In the music MOOC platform, when the number of times that two users visit item i reaches the set threshold, they are judged as neighboring users. In the big data platform, all items are classified according to similarity. According to the user's access to item i in the last time period, the content similar to item i is dynamically recommended for the user to access the platform in the next time period.

The predicted rating of item i by user u in the recommendation system is obtained by averaging the ratings of item i by the k neighboring users of user u . If a neighboring user of user u has not had any rating on item i , that user is removed as a neighboring user.

$$P_{u,i} = \frac{1}{k} \sum_{v \in N_u \cap S(i)} M_{v,i}, \quad (1)$$

where $S(i)$ denotes all users who have visited item i and $M_{v,i}$ denotes the predicted rating of item i by the user v .

$$P_{u,i} = \frac{\sum_{v \in N_u \cap S(i)} \text{sim}(u, v) M_{v,i}}{\sum_{v \in N_u \cap S(i)} |\text{sim}(u, v)|}, \quad (2)$$

where $\text{sim}(u, v)$ represents the similarity between the target user u and the user v in the adjacent user set N_u . Through the similarity calculation method, it is obtained that

$$P_{u,i} = \bar{M}_u \cdot \frac{\sum_{v \in N_u \cap S(i)} \text{sim}(u, v) \cdot (M_{v,i} - \bar{M}_v)}{\sum_{v \in N_u \cap S(i)} |\text{sim}(u, v)|}, \quad (3)$$

$$\text{sim}(i, j) = \frac{i \cdot j}{\|i\| \|j\|},$$

where \bar{M}_u and \bar{M}_v represent the average score of users u and v , respectively, and the calculation method is shown as follows:

$$\bar{M}_u = \frac{1}{N(I(u))} \sum_{i \in I_u} M_{ui}. \quad (4)$$

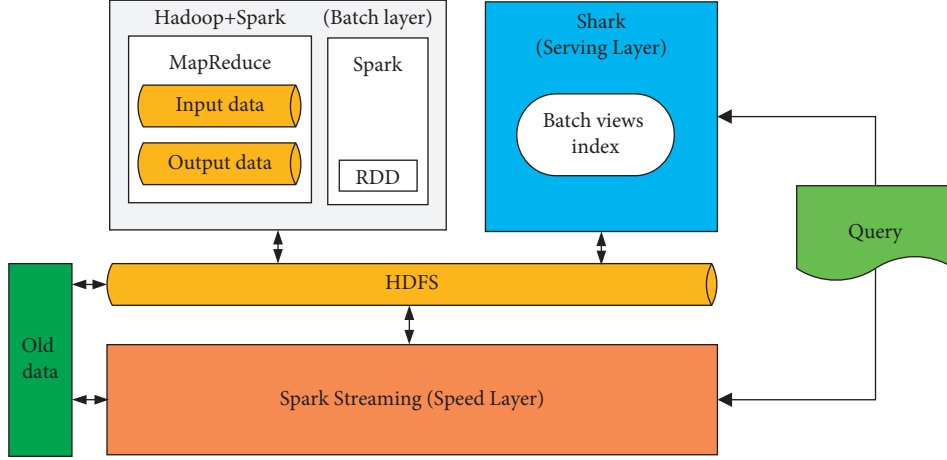


FIGURE 3: Distributed big data computing framework.

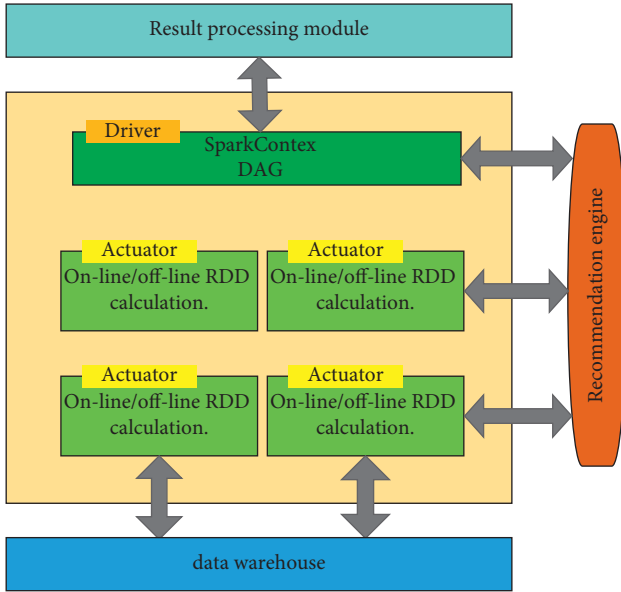


FIGURE 4: Spark platform architecture.

Take the top N music courses with higher prediction scores and recommend them to users. Let CF algorithm be recommended as T_1 and XGBoost model be recommended as T_2 , then the calculation method of the scoring result T of the mixed recommendation algorithm is as follows:

$$T = aT_1 + (1 - a)T_2. \quad (5)$$

In the evaluation of music course i by user u , let the weight of collaborative filtering algorithm be α_{ui} , the weight of recommendation algorithm of XGBoost model be β_{ui} , and the interference be e .

$$P_{ui} = \alpha_{ui}M_{ui} + \beta_{ui}M'_{ui} + e. \quad (6)$$

The dataset consisting of α_{ui} and β_{ui} is sparse, and most of the data values are 0. In the process of data fitting by algorithm, α_{ui} and β_{ui} can be written as α_u and β_u . Therefore, formula (6) can be written as

$$P_{ui} = \alpha_u M_{ui} + \beta_u M'_{ui} + e. \quad (7)$$

According to formula (5), we can know that $\alpha_u + \beta_u = 1$.

Finally, the highest value of personalized recommendation score of dynamic users is solved, which is transformed into the minimum difference between the actual score value and the predicted score mean. The calculation method is shown as follows:

$$\min_{\alpha_u, \beta_u} = \sum_{R \in S} (M_{u,j} - \bar{M}_u)^2. \quad (8)$$

There are four main indexes to measure the recommendation algorithm, namely, precision, recall, mean absolute error (MAE), and root mean square error (RMSE):

$$\text{precision} = \frac{|S_u \cap S'_u|}{|S'_u|}, \quad (9)$$

$$\text{recall} = \frac{|S_u \cap S'_u|}{|S_u|},$$

where S_u denotes the set of music courses with the highest predicted values using the mixed algorithm and S'_u denotes the set of music courses with the highest ratings for user u in the training sample.

$$\text{MAE} = \frac{1}{|\text{count}(S)|} \sum_{R \in S} |R_{u,j} - M_{u,j}|,$$

$$\text{RMSE} = \sqrt{\frac{\sum_{R \in S} (R_{u,j} - M_{u,j})^2}{\text{count}(S)}}, \quad (10)$$

where $R_{u,j}$ denotes the actual rating value of music course j by user u , $M_{u,j}$ denotes the rating value of music course j by user u after using the recommendation algorithm, and $\text{count}(S)$ denotes the total number of user rating sets. In this paper, MAE and RMSE are used as the main metrics for recommendation performance evaluation.

3.2. Determination of Collaborative Weight Factor. In this paper, shuffled frog leading algorithm [26] is used to optimize the weight factor. Initialize the weight factor as frog swarm, and then perform shuffled frog leading algorithm to solve the weight factor value. In the process of weight factor optimization, the reciprocal of the difference between the predicted value and the actual value is used as the fitness function of the mixed shuffled frog leading algorithm. In the $(t + 1)$ -th calculation iteration, the results after the t -th iteration are used. The maximum of the frog fitness function is $X_b(t)$, and the minimum of the frog fitness function is $X_w(t)$. In order to ensure that the frogs in the group can move closer to the frog with the largest fitness function value, the frog with the smallest fitness function value starts to move continuously, and the moving method is as follows:

$$\Delta_w(t) = \text{rand}(X_b(t) - x(t)). \quad (11)$$

$$X_w(t + 1) = X_w(t) + \Delta_w(t), R_{\min} \leq \Delta_w(t) \leq R_{\max}. \quad (12)$$

If the value of $X_w(t + 1)$ solved at time $t+1$ is larger than $X_w(t)$, that is, it has better fitness, then replace $X_w(t)$ with $X_w(t + 1)$. Otherwise, continue to execute formulas (11) and (12). With regard to frog moving step size, step size factor C can be introduced, so the formula for calculating the i -th moving distance of the k -th frog is as follows:

$$d_i = \text{rand} \times (X_b^k - X_w^k) \times C, \quad (13)$$

$$C = C_{\min} + \frac{i_{\text{now}}}{G_{\text{global}}} \times (C_{\max} - C_{\min}),$$

where C_{\min} and C_{\max} are, respectively, the minimum and maximum moving steps of frogs in the current group, which can be set according to the actual situation. G_{global} is the sum of the fitness values of all the frogs in the group. i_{now} is the number of times the frog moves at the current time.

When the fitness values of all frogs in the group are close to $X_b(t)$ and the error is within the set threshold, the algorithm iteration stops. Output the frog distribution map at the current moment, which is the optimal solution.

4. Experiment and Result Analysis

4.1. Experimental Environment and Dataset. This system runs on a Spark cluster with 6 nodes, and the configuration of each node is exactly the same. The operating system of each node is CentOS 6.5 Linux, 32 GB RAM, 2T hard disk, E5645 @ 2.40 GHz 6-Core CPU, Spark version 1.5.1, and 1000 Mbps network card. The main hardware configuration information and software configuration information of the experiment are shown in Tables 1 and 2, respectively.

This paper adopts the dataset of the <http://www.cmooc.com> website, which mainly includes two datasets: user's course selection record and user's operation log. By analyzing the user's course selection records, we can obtain that the number of users is 120069 and the number of courses is 68, and we can analyze the interaction between users and courses. We can find that, among the total 18 music courses, only a few of the three courses have a large

number of students, about 10,000. However, most music courses have a small number of students, and only about 2,000 users attend these courses. It is not difficult to find that the vast majority of users only choose one music course, so the users' course selection is very sparse, and the dataset has obvious cold start problem. In order to facilitate the experiment and avoid the cold start problem, we set the threshold of cold start to 6.

4.2. Acceleration Ratio. In order to verify the influence of the Spark platform on the speed of music MOOC resources, the acceleration ratio of Spark recommendation relative to single-machine recommendation is solved.

$$S = \frac{T_a}{T_s}, \quad (14)$$

where T_a and T_s are recommended times for single machine and Spark multinode, respectively.

From the Spark acceleration performance shown in Table 3, it can be seen that when the number of working nodes increases, the Spark acceleration effect becomes more obvious. The larger the sample size, the more significant the influence of the number of working nodes on the speedup ratio. The sample capacity of Data 1 is 12.95 MB, and when the number of working nodes reaches 6, the speedup ratio is only 0.001 higher than that of the single machine. However, when the sample size is 8.03 GB, the speedup ratio increases by 42.907 compared with the single machine, so the Spark platform improves the recommendation efficiency of large-capacity samples and is especially suitable for the recommendation task of large-scale data.

4.3. Recommended Speed and Accuracy Analysis. The recommendation time of the recommendation algorithm proposed in this paper is compared between MapReduce architecture and Spark architecture. As shown in Figure 5, it can be seen that Spark architecture takes less time to obtain Top-N recommendation [27] (MOOC recommendation belongs to Top-N recommendation) than MapReduce architecture.

In order to verify the effectiveness of the proposed mixed collaborative filtering recommendation algorithm, it is compared with traditional collaborative filtering recommendation [28], recommendation based on the XGBoost model [21], and recommendation based on Canopy clustering [29]. Each algorithm was randomly repeated 30 times, and the average values of MAE and RMSE were calculated. The comparison results of MAE with different sparsity are shown in Table 4 and Figure 6. The comparative results of RMSE with different sparsity are shown in Table 5 and Figure 7.

From the above experimental results, it can be seen that, with the increasing sparsity, the prediction accuracy of various collaborative filtering recommendations is also increasing. It shows that the realization of these four algorithms is trained under the condition of a large amount of data. If a user's associated data under the MOOC platform is

TABLE 1: Hardware configuration.

Number	Node name	CPU	Internal storage capacity	Hard disc capacity (TB)
1	Master	E5645 @ 2.40 GHz 6-Core	32 GB RAM	2
2	Slave01	E5645 @ 2.40 GHz 6-Core	32 GB RAM	2
3	Slave02	E5645 @ 2.40 GHz 6-Core	32 GB RAM	2
4	Slave03	E5645 @ 2.40 GHz 6-Core	32 GB RAM	2
5	Slave04	E5645 @ 2.40 GHz 6-Core	32 GB RAM	2
6	Slave05	E5645 @ 2.40 GHz 6-Core	32 GB RAM	2

TABLE 2: Software configuration.

Number	Node name	Operating system	Spark's version
1	Master		
2	Slave01		
3	Slave02		
4	Slave03	CentOS 6.5 Linux	Spark 1.5.1
5	Slave04		
6	Slave05		

TABLE 3: Spark acceleration performance.

Sample sets	Sample set size	Number of nodes involved in the calculation	Acceleration ratio
Data 1	12.95 MB	1	1.000
		3	1.001
		6	1.001
Data 2	656.82 MB	1	1.000
		3	1.013
		6	1.024
Data 3	1.42 GB	1	1.000
		3	6.473
		6	13.139
Data 4	8.03 GB	1	1.000
		3	22.287
		6	43.907

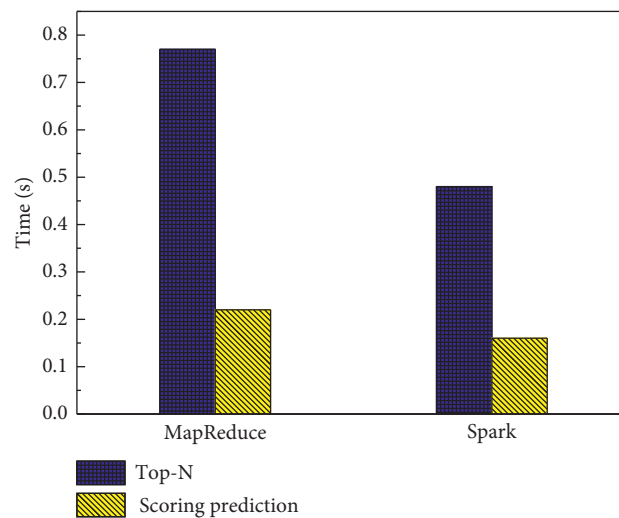


FIGURE 5: Recommended time comparison.

TABLE 4: Comparison results of MAE with different sparsity.

	Sparsity		
	10%	20%	30%
Traditional collaborative filtering	0.594	0.450	0.416
XGBoost	0.582	0.443	0.411
Canopy cluster	0.542	0.421	0.409
Mixed collaborative filtering	0.538	0.418	0.406

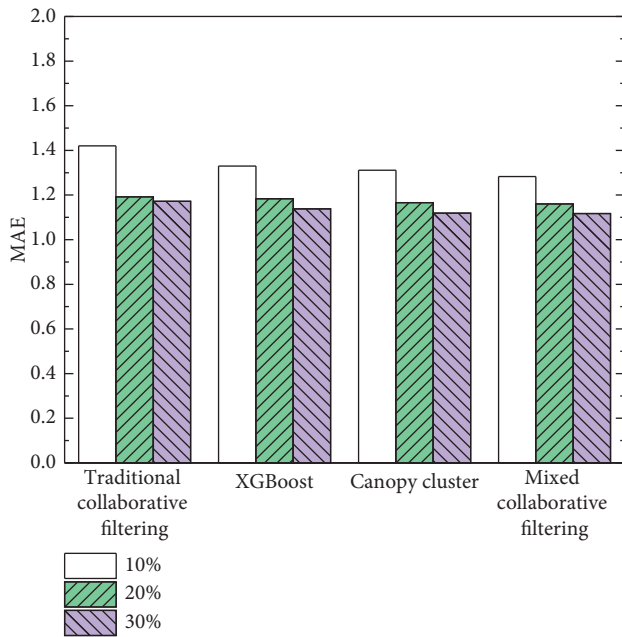


FIGURE 6: MAE comparison results at different sparsity.

TABLE 5: Comparison results of RMSE with different sparsity.

	Sparsity		
	10%	20%	30%
Traditional collaborative filtering	1.420	1.192	1.172
XGBoost	1.330	1.183	1.138
Canopy cluster	1.311	1.166	1.119
Mixed collaborative filtering	1.283	1.160	1.117

less (sparsity is lower), the effective recommendation of the user's personalized data cannot be completed at this time, which accords with the characteristics of data mining. When there are few user-related feature data, it is difficult to mine valuable data for the user, and it is even more difficult to achieve effective recommendations according to the user's preferences and habits. Under the same sparsity condition, compared with other existing recommendation methods, the MAE performance of the proposed mixed clustering collaborative filtering recommendation is better. The results show that the proposed method significantly improves the accuracy of personalized service recommendation of music MOOC.

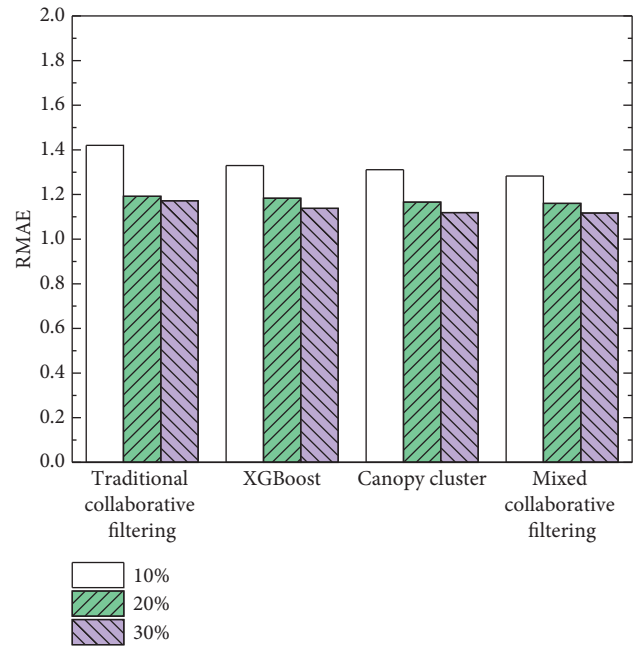


FIGURE 7: RMSE comparison results at different sparsity.

5. Conclusions

Personalized recommendation service of music MOOC resources enhances users' experience, and users can efficiently obtain effective data of the platform, avoiding searching and searching in massive data. The good experience of personalized service is based on an accurate recommendation algorithm. If the applicability of the recommendation algorithm is not strong, and the service recommended to users is irrelevant to users, it will burden users with junk data and reduce their experience. In this paper, the XGBoost algorithm and CF algorithm are combined to improve the performance of personalized recommendation on the Spark platform. Experiments show that the proposed algorithm has an excellent performance in accuracy and real time and has strong popularization and application value. Further research will be carried out on the problem of excessive differences in sub-datasets caused by randomly dividing datasets.

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 author declares that there are no conflicts of interest to report regarding the present study.

References

- [1] T. He, "Information overload and interaction overload as two separate attributes of information overload syndrome," *Journal of Enterprising Culture*, vol. 28, no. 03, pp. 263–279, 2020.
- [2] C. Gunaratne, N. Baral, and W. Rand, "The effects of information overload on online conversation dynamics,"

- Computational & Mathematical Organization Theory*, vol. 26, no. 5, pp. 1–22, 2020.
- [3] O. Amir, B. J. Grosz, K. Z. Gajos, and L. Gultchin, “Personalized change awareness: reducing information overload in loosely-coupled teamwork,” *Artificial Intelligence*, vol. 275, no. 10, pp. 204–233, 2019.
 - [4] H. Lee and K. Choi, “Recommending valuable ideas in an open innovation community: a text mining approach to information overload problem,” *Industrial Management & Data Systems*, vol. 101, no. 6, pp. 158–168, 2018.
 - [5] E. Zkan and M. Tolon, “The effects of information overload on consumer confusion: an examination on user generated content,” *Social Science Electronic Publishing*, vol. 60, no. 5, pp. 4473–4482, 2018.
 - [6] J. Cheng, F. Zhang, and X. Guo, “A syntax-augmented and headline-aware neural text summarization method,” *IEEE Access*, vol. 8, Article ID 218360, 2020.
 - [7] H. Qian, “Automatic generation of MOOC course notes based on UGC knowledge fusion,” *Information Studies: Theory & Application*, vol. 39, no. 2, pp. 272–286, 2019.
 - [8] M. A. Kaufhold, M. Bayer, and C. Reuter, “Rapid relevance classification of social media posts in disasters and emergencies: a system and evaluation featuring active, incremental and online learning,” *Information Processing & Management*, vol. 57, no. 1, pp. 1–12, 2020.
 - [9] J. Tieman, L. Miller-Lewis, D. Rawlings, D. Parker, and C. Sanderson, “The contribution of a MOOC to community discussions around death and dying,” *BMC Palliative Care*, vol. 17, no. 1, pp. 31–39, 2018.
 - [10] G. Chen, D. Davis, M. Krause, E. Aivaloglou, C. Hauff, and G.-J. Houben, “From learners to earners: enabling MOOC learners to apply their skills and earn money in an online market place,” *IEEE Transactions on Learning Technologies*, vol. 11, no. 2, pp. 264–274, 2018.
 - [11] R. F. Babiceanu and R. Seker, “Big Data and virtualization for manufacturing cyber-physical systems: a survey of the current status and future outlook,” *Computers in Industry*, vol. 81, pp. 128–137, 2016.
 - [12] M. Aparicio, T. Oliveira, F. Bacao, and M. Painho, “Gamification: a key determinant of massive open online course (MOOC) success,” *Information & Management*, vol. 56, no. 1, pp. 39–54, 2019.
 - [13] X. Peng and Q. Xu, “Investigating learners’ behaviors and discourse content in MOOC course reviews,” *Computers & Education*, vol. 143, no. 1, p. 1, Article ID 103673.14, 2020.
 - [14] Y. D. Min, C. J. Bonk, and J. Kim, “An investigation of under-represented MOOC populations: motivation, self-regulation and grit among 2-year college students in Korea,” *Journal of Computing in Higher Education*, vol. 12, pp. 1–22, 2021.
 - [15] P. D. Barba, D. Malekian, and E. A. Oliveira, “The importance and meaning of session behaviour in a MOOC,” *Computers & Education*, vol. 146, no. 5, pp. 1–18, 2020.
 - [16] L. Yao, Q. Z. Sheng, A. H. H. Ngu, J. Yu, and A. Segev, “Unified collaborative and content-based Web service recommendation,” *IEEE Transactions on Services Computing*, vol. 8, no. 3, pp. 453–466, 2015.
 - [17] G. Zheng, H. Yu, and W. Xu, “Collaborative filtering recommendation algorithm with item label features,” *International Core Journal of Engineering*, vol. 6, no. 1, pp. 160–170, 2020.
 - [18] X. Deng, T. Liu, W. Li, F. Liu, and J. Peng, “A latent factor model of fusing social regularization term and item regularization term,” *Physica A: Statistical Mechanics and Its Applications*, vol. 525, pp. 1330–1342, 2019.
 - [19] Y. Zhang, Z. Liu, and C. Sang, “Unifying paragraph embeddings and neural collaborative filtering for hybrid recommendation,” *Applied Soft Computing*, vol. 106, no. 6, pp. 107–116, 2021.
 - [20] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, “Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency,” *Knowledge and Information Systems*, vol. 54, no. 2, pp. 1–32, 2017.
 - [21] C. Wang, C. Deng, and S. Wang, “Imbalance-XGBoost: leveraging weighted and focal losses for binary label-imbalance classification with XGBoost,” *Pattern Recognition Letters*, vol. 136, no. 12, pp. 353–363, 2020.
 - [22] S. Babu, R. Pitchai, and S. Anjanayya, “Web log analysis using Spark with solution recommendation,” *Materials Today Proceedings*, no. 3, pp. 1235–1241, 2021.
 - [23] P. S. Sundari and M. Subaji, “Integrating sentiment analysis on hybrid collaborative filtering method in a big data environment,” *International Journal of Information Technology and Decision Making*, vol. 19, pp. 287–296, 2020.
 - [24] S. Kundu and U. Maulik, “Cloud deployment of game theoretic categorical clustering using Apache spark: an application to car recommendation,” *Machine Learning with Applications*, vol. 2, no. 12, pp. 1–22, 2021.
 - [25] W. Xue and T. Wu, “Active learning-based XGBoost for cyber physical system Against generic AC false data injection attacks,” *IEEE Access*, vol. 8, Article ID 144575, 2020.
 - [26] S. Anandamurugan and T. Abirami, “Antipredator adaptation shuffled frog leap algorithm to improve network life time in wireless sensor network,” *Wireless Personal Communications*, vol. 94, no. 4, pp. 1–12, 2017.
 - [27] S. Gopalani and R. Arora, “Comparing Apache spark and map reduce with performance analysis using K-means,” *International Journal of Computer Application*, vol. 113, no. 1, pp. 8–11, 2015.
 - [28] Y. Cai, H.-f. Leung, Q. Li, H. Min, J. Tang, and J. Li, “Typicality-based collaborative filtering recommendation,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 3, pp. 766–779, 2014.
 - [29] Y. Yang, H. Wu, and C. Yan, “Medical consumable usage control based on Canopy_K-means clustering and WARM,” *Journal of Combinatorial Optimization*, vol. 22, no. 8, pp. 68–77, 2019.

Research Article

Research on Ecoenvironmental Quality Evaluation System Based on Big Data Analysis

Pingheng Li 

Business School, Huanggang Normal University, Huanggang, Hubei 438000, China

Correspondence should be addressed to Pingheng Li; lipingheng@hgnu.edu.cn

Received 24 December 2021; Revised 19 January 2022; Accepted 28 January 2022; Published 7 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Pingheng 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.

Comprehensive and objective evaluation of ecological environment quality is of great significance to regional sustainable development. In this study, Landsat remote sensing images of 1991, 2000, 2004, 2010, 2013, 2018, and 2019 are selected to evaluate the changes of ecological environment quality in the Headwaters of Dongjiangyuan River by using remote sensing ecological index RSEI. The influencing factors of ecological environment change in Dongjiangyuan River are also discussed. The results showed that, from 1991 to 2019, the ecoenvironmental quality of the Dongjiangyuan River showed a good trend of development. Humidity index, greenness index, and dryness index all fluctuated in a small range; the greenness and dryness showed an overall increase. The average temperature in the Headwaters of the Dongjiangyuan River presents a rising trend. This study establishes the evaluation system of ecological environment quality from two dimensions of time and space and gives the change rule of environmental quality quantitatively, which provides the theoretical basis for the ecological environment management of Dongjiangyuan River.

1. Introduction

The ecological environment is a compound ecosystem that affects human life and production, which plays an extremely important role. Ecological environment determines the comfort level of human life. Meanwhile, the quality of ecological environment also restricts the development of social economy. In recent years, with the expansion of urban scale, that human beings have damaged the ecological environment increasingly, which results in a sharp reduction of vegetation, soil erosion, reduction in biodiversity, and other negative effects [1, 2]. Therefore, to accurately assess the contradiction between the needs of human development and the ecological environment, to establish an effective quality evaluation system, and to complete the systematic evaluation of the ecological environment for specific regions are hot issues that needed to be solved urgently by all countries in the world [3].

The traditional ecological environment assessment mainly combines survey data for statistical analysis, which has the disadvantages of high cost, being time-consuming

and laborious, and subjective interference [4]. The emergence of big data solves this problem. The big data refers to the collection of data that cannot be captured, managed, and processed by conventional software tools within a certain period of time. The high-growth and diversified information asset requires a new processing mode to have stronger decision-making ability, insight, and discovery ability and process optimization ability. The big data is different from the concepts of “massive data” and “superlarge scale data” fundamentally. In addition to considering the data capacity, the big data focuses more on the diversity of data types, the efficiency acquisition speed of data, the variability, authenticity, complexity, and value of data. Pioneering abandons the excessive desire for causality in traditional research methods and pays more attention to correlation [5, 6]. The quality assessment of ecological environmental that is based on the analysis of the large data can do acquisition, processing, analysis, and application of all kinds of spatial data and nonspatial data, such as POI data, track data, and mobile phone signal data as well as the public comments on open source data, by intelligent means, tools, or software

in the limited time, compared with the traditional pattern of “data” analysis and evaluation. The statistical analysis that is based on big data can make a comprehensive evaluation of ecological environment quality objectively and quantitatively [7].

The foundation of ecological quality evaluation system based on big data is the way of acquiring big data. With the extensive development of remote sensing technology, it has become an effective way to acquire big data of surface resources by means of satellite images. The data from remote sensing can obtain the distribution of ground facilities, environmental pollution, and other problems quickly and comprehensively, combined with the corresponding analysis means, which can achieve a comprehensive and objective evaluation of ecological environment quality [8, 9]. Based on this, it has great value for research to establish an ecological environment assessment model to systematically evaluate the ecological environment quality of a specific region by using remote sensing to obtain big data [10].

2. Related Work

The evaluation of ecological environment quality overseas pays more attention to the practicability and maneuverability of process and result. Paula et al. [11] selected the indexes from the perspective of land suitability, and the index system including natural environment suitability, biological environment suitability, and functional suitability was constructed. The suggestions based on the analysis results for promoting urban sustainable development were put forward. Alatang et al. [12] studied the relationship between urban economic development and environmental quality of 43 countries by using quantitative statistical calculation method and established the well-known environmental Kuznets curve hypothesis. By Richard et al. [13] based on the perspective of ecological security, combining with three subsystems of forest, water area, and grassland, the ecological security evaluation index system was constructed by selecting 11 indexes, and the ecological security of the area near the Colorado River was evaluated scientifically and comprehensively by cluster analysis method. By Valentina et al. [14], using analytic hierarchy process (AHP), an index system with 11 indexes was constructed according to biomolecules and physicochemical elements to evaluate the ecological environment quality in three semiclosed coastal areas. Muhammad et al. [15] constructed the Mediterranean submarine cave ecosystem which uses the theoretical model of structure operation, and the quality index framework (EBQI) was established based on the structural operation theory model, to evaluate the ecological environmental quality of Mediterranean submarine cave and its surrounding environment.

The evaluation of ecological environment quality in China has gone through a long process. From the definition and principle level, it has gradually developed to qualitative and quantitative evaluation research on ecological environment by using mathematical and physical methods. The ecological system evaluation in China is gradually becoming institutionalized and standardized, and the weight treatment

of evaluation elements and evaluation content are increasingly enriched [16]. Zhang et al. [17] used the remote sensing images of Tai'an city in two periods; the EI index was graded and evaluated. The results showed that the ecological environment quality was mainly good and distributed in southwest China. Xu et al. [18] selected greenness, humidity, heat, and dryness indexes, and the principal component analysis was used to build an ecological environment quality evaluation system. The proposal and application of the new Remote Sensing Based Ecology Index (RSEI) broadened the research direction of many scholars. Jia et al. [19] established the evaluation system from three aspects: ecological environment level, ecological environment pressure, and ecological environment protection. The index weight was assigned by entropy weight method and the ecological environment quality of Heilongjiang province was dynamically evaluated by comprehensive index method. Huang et al. [20] constructed the fuzzy comprehensive evaluation index system of ecosystem for the region of Yiqiao mining area, Shandong Province, the membership function is obtained from the original data, and the improvement of entropy technology is used to assign weights to indicators, and the ecological health of the mining area is quantitatively studied. Liu et al. [21] chose Bashang Plateau ecological area as the research object, the ecological environment quality, ecological environment quality in mountain ecological area, and ecological environment quality of plain ecological area as the evaluation index system of target layer and selected 6 ecological environmental quality factors that include the ecosystem, the natural resources, the biological diversity, the biological disasters, the food security, and the social ecological system, which constitutes system layer, with the analytic hierarchy process (ahp) to evaluate the index system. Gao et al. [22] used sustainable development theory and ecological economics principle; 12 indexes were selected, from the angle of economy, environment, and society, that constructed the evaluation index system of ecological environment quality of Xi'an city. The analytic hierarchy process and fuzzy comprehensive evaluation are used to evaluate it quantitatively. Zhou et al. [23] selected the factors that lead to ecological fragility. The ecological environment quality of Beipanjiang river basin was analyzed comprehensively and systematically. The law of ecological development in this area was revealed, and there is still a lot of management space in environmental protection. Li et al. [24] took 10 years as the research period; the ecological environment of Chaohu Lake basin was analyzed. The results showed that vegetation coverage and dryness had a greater effect on the ecological environment in this region.

Principal component analysis: Xu et al. [25] calculated and analyzed the data in the index system by principal component analysis model and made a comprehensive evaluation and comparative study on the ecological environmental quality level of all cities in Anhui Province, providing valuable basic data for the development of ecological civilized cities and future environmental planning in Anhui Province. Cheng and Chi [26] in 2011 used the DPSIR framework to build an index system and established an evaluation model based on the nuclear principal component

analysis method. Ten representative cities were selected for evaluation. Sun [27] in 2014 constructed the urban ecological environment quality of Nanjing from the two perspectives of environmental pollution and ecological damage. Principal component analysis was chosen as the evaluation method, and the evaluation object was the ecological environment quality of Nanjing in five years. Chun [28] in 2015, based on the basis of a comprehensive evaluation, according to the specific region characteristic and the development of the Xi'an in Shanxi Province, established a comprehensive evaluation index system of eco-city development, using the method of principal component analysis on the comprehensive development of urban ecological quality, and evaluated the quality of each subsystem analysis, according to the result of evaluation data to find out the reason. The evaluation of the coordinated development of each subsystem reflects the current situation of ecological environment quality in Xi'an. Fuzzy comprehensive evaluation method: In the evaluation of environmental quality, Wang [29] in 2017 selected 40 indicators to construct an index system based on the characteristics of the county (city) development from five perspectives of human settlement environment, social equity, industrial structure, ecological construction, and resource utilization and evaluated and analyzed the ecological environment quality of county (city) with fuzzy comprehensive analysis method.

To sum up, a lot of evaluation and analysis work were done by experts and scholars at home and abroad that built ecoenvironmental quality evaluation systems based on different evaluation objects. However, the single evaluation on the space or time of ecological environment was carried out by existing literature, which led to one-sided analysis and conclusion. In this article, with the east of river basin as the research object, the remote sensing and geographic information technology monitoring were used to monitor the ecological and environmental quality of the watershed. The RSEI model was selected to quantitatively study the ecological environment quality of Dongjiangyuan River watershed. The spatial distribution and temporal variation of ecological environment quality in the source basin of Dongjiangyuan River were analyzed. It provides a scientific basis for future environmental management and ecological civilization construction in Dongjiangyuan River basin.

3. The Principle and Analysis Method of the Ecoenvironmental Quality Assessment System

In order to evaluate regional ecological quality quickly, Xu proposed the remote sensing ecological index (RSEI). The remote sensing ecological index model is based on remote sensing images to extract greenness, humidity, dryness, and

heat index, through principal component analysis; the four factors are integrated to monitor the ecological environment and evaluate the ecological environment quality.

3.1. The Principles of RSEI Model

3.1.1. The Humidity Indicators. The humidity index is closely related to the moisture content of vegetation and soil, which is widely used in ecological environment monitoring and evaluation. The range of wet values is between $[-1, 1]$, and the larger the value is, the higher the humidity is. The humidity index can be represented by the wet component in the $k-T$ variation. Different formulas used by Landsat sensors are as follows:

TM data:

$$WET = 0.0315\rho_{blue} + 0.2021\rho_{green} + 0.3102\rho_{red} + 0.1594\rho_{nir} - 0.6806\rho_{swir1} - 0.6109\rho_{swir2}. \quad (1)$$

OLI data:

$$WET = 0.1511\rho_{blue} + 0.1973\rho_{green} + 0.3283\rho_{red} + 0.3407\rho_{nir} - 0.7171\rho_{swir1} - 0.4559\rho_{swir2}, \quad (2)$$

where ρ_{blue} is the blue bands, ρ_{green} is the green bands, ρ_{red} is the red bands, ρ_{nir} is the near-infrared band, ρ_{swir1} is the reflectivity of shortwave infrared 1, and ρ_{swir2} is the reflectivity of shortwave infrared 2.

3.1.2. The Green Degree Index (NDVI). Normalized Difference Vegetation Index (NDVI) is the most widely used vegetation index, which can effectively reflect the growth status of plants. The range of NDVI value is between $[-1, 1]$, and the larger the value is, the higher the vegetation coverage is. Therefore, normalized vegetation index (NDVI) was used to represent the greenness index. The formula is as follows:

$$NDVI = \frac{(\rho_{nir} - \rho_{red})}{(\rho_{nir} + \rho_{red})}, \quad (3)$$

where ρ_{nir} and ρ_{red} are the near-infrared reflectance and the reflectivity of the red band.

3.1.3. The Dryness Index (NDSI). The soil drying represents the degree of land exposure and dryness. The continuous desiccation of soil has a serious impact on the ecological environment quality in this region. It is also one of the important factors of ecosystem imbalance. In this paper, the average value of the building index (IBI) and bare soil index (SI) was used to construct the dryness index (NDSI). The NDSI value ranges in $[-1, 1]$, and the higher the value is, the higher the degree of drying is. The formula is as follows:

$$\begin{aligned}
IBI &= \frac{2\rho_{\text{swir1}}/(\rho_{\text{swir1}} + \rho_{\text{nir}}) - \rho_{\text{nir}}/(\rho_{\text{nir}} + \rho_{\text{red}}) - \rho_{\text{green}}(\rho_{\text{green}} + \rho_{\text{swir1}})}{2\rho_{\text{swir1}}/(\rho_{\text{swir1}} + \rho_{\text{nir}}) + \rho_{\text{nir}}/(\rho_{\text{nir}} + \rho_{\text{red}}) + \rho_{\text{green}}(\rho_{\text{green}} + \rho_{\text{swir1}})}, \\
SI &= \frac{(\rho_{\text{swir1}} + \rho_{\text{red}}) - (\rho_{\text{nir}} + \rho_{\text{blue}})}{(\rho_{\text{swir1}} + \rho_{\text{red}}) + (\rho_{\text{nir}} + \rho_{\text{blue}})}, \\
NDSI &= \frac{(IBI + SI)}{2}.
\end{aligned} \tag{4}$$

3.1.4. *The Heat Index (LST)*. The heat indicators are expressed in terms of surface temperature. The surface temperature is the temperature of the ground that absorbs solar heat radiation, which affects the growth and development of vegetation and has a strong intervention effect on the water cycle. At the same time, it is also one of the factors affecting the evaporation and transpiration of natural water and indirectly affects the change of ecological environment. The larger the VALUE of LST, the higher the temperature. The surface temperature in this paper is obtained by modifying the brightness temperature, and the formula is as follows:

$$\begin{aligned}
L_{\text{TIR}} &= \text{gain} \times \text{DN} + \text{bias}, \\
T &= \frac{K_2}{\ln(K_1/L_{\text{TIR}} + 1)},
\end{aligned} \tag{5}$$

where L_{TIR} is radiation calibration of the thermal infrared band; K_1, K_2 are calibration coefficient.

$$\text{LST} = \frac{T}{[1 + (\lambda T/\rho) \ln \varepsilon]}, \tag{6}$$

where $T, \lambda,$ and ε are the brightness temperature, the central wavelength of the thermal infrared band, and the surface emissivity.

TM data: $K_1 = 607.76 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \cdot \mu\text{m}^{-1}, K_2 = 1260.56 \text{ K},$
 $\text{gain} = 0.055, \text{ bias} = 1.18243, \lambda = 11.45 \mu\text{m}, \rho = 1.438 \times 10^{-2} \text{ m} \cdot \text{K}.$

Landsat8 data: $K_1 = 774.89 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \cdot \mu\text{m}^{-1},$
 $K_2 = 1321.08 \text{ K}, \text{ gain} = 3.342, \text{ bias} = 0.1, \lambda = 10.90 \mu\text{m}, \rho = 1.438 \times 10^{-2} \text{ m} \cdot \text{K}.$

The land surface emissivity is calculated by the NDVI threshold method proposed by Sobrino, and the formula is as follows:

$$\begin{aligned}
\varepsilon &= 0.004f + 0.986, \\
f &= \frac{(\text{NDVI} - \text{NDVI}_{\text{soil}})}{(\text{NDVI}_{\text{veg}} - \text{NDVI}_{\text{soil}})},
\end{aligned} \tag{7}$$

where f is the vegetation coverage, $\text{NDVI}_{\text{soil}}$ is the NDVI value of bare soil or no vegetation-covered area, and NDVI_{veg} is the NDVI value of complete vegetation coverage.

3.2. The Principal Component Analysis

3.2.1. *The Basic Principles of Principal Component Analysis*. The principal component analysis (PCA) was proposed by Karl Pearson, a British mathematician, in 1901, to realize multidimensional data compression. The principal component analysis is a method used in mathematical statistics, mathematical analysis, and mathematical modeling to transform multiple variables into a few variables through mathematical transformation. The principal component analysis is to obtain another set of unrelated variables through matrix transformation, which is the idea of data dimension reduction. The principal component analysis converts complex elements into N principal components when introducing multiple variables, to simplify the problem and achieve scientific and effective results.

3.2.2. *The Mathematical Model of Principal Component Analysis*. From the basic principle of principal component analysis, it can be known that principal component analysis is an idea of data dimension reduction, which obtains a new set of independent variables with certain correlations through mathematical transformation.

The principal component analysis is described in a mathematical way. For a data set X with n samples, $X_1, X_2 \dots X_p,$ and P variables, the data matrix is

$$\begin{aligned}
X &= \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \\
&= [x_1, x_2, \dots, x_p],
\end{aligned} \tag{8}$$

where $x_i = (x_{1i}, x_{2i}, \dots, x_{ni})^T, i = 1, 2, \dots, p.$

The principal component analysis is to synthesize the original P observation $X_1, X_2, \dots, X_p,$ with variables to form P new variables.

$$\begin{bmatrix} F_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p \\ F_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p \\ \dots \\ F_p = a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pp}x_p \end{bmatrix}. \tag{9}$$

It can be written as theta:

$$F_1 = w_{1i}x_{i1} + w_{2i}x_{i2} + \cdots + w_{pi}x_{ip}, \quad i = 1, 2, \dots, p, \quad (10)$$

where x_i , F_i are all n -dimensional vectors

The above model should satisfy the following three conditions:

- (1) There is no correlation between F_i and F_j ($i \neq j$, $i, j = 1, 2, \dots, p$)
- (2) The variance of F_1 is greater than the variance of F_2 , and F_2 is greater than the variance of F_3 , and so on
- (3) $w_{k1}^2 + w_{k2}^2 + \cdots + w_{kp}^2 = 1$, $k = 1, 2, \dots, p$

If all three conditions are satisfied, the transformation results in new independent variables.

3.3. The RSEI Model Calculation Based on Principal Component Analysis. In this study, the principal component analysis was used to integrate greenness, humidity, dryness, and heat, so as to achieve the purpose of expressing information with a single index. Since the dimensions of all indicators are not uniform, normalization of all indicators is required before principal component analysis to make the range of index values $[0, 1]$. The normalization formula is as follows:

$$N = \frac{I - I_{\min}}{I_{\max} - I_{\min}}, \quad (11)$$

where N is the normalized pixel value, I_i is the original value of pixel i , I_{\max} is the maximum value of pixels, and I_{\min} is the minimum value of pixels.

After the normalization of humidity, greenness, dryness, and heat indexes, the four indexes are superimposed to synthesize a new layer. Then apply the principal component transformation to the new layer. If ER-Mapper is used for principal component analysis, RSEI₀ is obtained by using the formula of the first principal component, and then the remote sensing ecological index is obtained by normalization.

$$\begin{aligned} \text{RSEI}_0 &= 1 - \{\text{PCI}[f(\text{Wet}, \text{NDVI}, \text{LST}, \text{NDSI})]\}, \\ \text{RSEI} &= \frac{(\text{RSEI}_0 - \text{RSEI}_{0\min})}{(\text{RSEI}_{0\max} - \text{RSEI}_{0\min})}, \end{aligned} \quad (12)$$

where RSEI is the remote sensing ecological index, which ranges in $[0, 1]$. The higher the RSEI value is, the better the ecological environment quality is. Conversely, the smaller the RSEI value is, the worse the quality of the ecological environment is.

The ENVI software is used for principal component analysis; the larger the first principal component data value obtained, the better the ecological environment quality. Therefore, the PCI can be normalized directly to generate remote sensing ecological index. The formula is as follows:

$$\text{RSEI} = \frac{(\text{PCI} - \text{PCI}_{\min})}{(\text{PCI}_{\max} - \text{PCI}_{\min})}, \quad (13)$$

where the PCI_{\min} is the minimum value of the first principal component; the PCI_{\max} is the maximum value of the first principal component

4. Results and Analysis of Ecological Environment Assessment in Dongjiangyuan River

4.1. The Data Source. The original data used in this paper are mainly Landsat series remote sensing images and DEM numerical elevation data, combined with the 1:100,000 land-use status map of the study area in 2015 and other basic data, mainly including social and economic data and natural resource data of the study area. For example, soil conditions, vegetation types, per capita GDP, and other data are shown in Tables 1 and 2.

All Landsat series remote sensing data used in this paper came from USGS website. In this paper, three Landsat remote sensing images were selected, namely, September 6, 1998 (Landsat-5 TM), September 1, 2008 (Landsat-5 TM), and August 9, 2017 (Landsat-8 OLI). The remote sensing image data of the three phases were in the same season. The unity of time is fully considered in the selection, and the interpretation results have good space-time contrast, which can meet the needs of this study.

In the process of further index screening, rough set equivalence relation in quantitative analysis method was used to screen the index, and after the index screening, expert consultation method was used to further improve the index system of ecological environment quality in Yunnan Province. By using the combination of interval hesitation fuzzy set and entropy weight method, the index weight could be obtained more accurately by scoring from several experts. Data processing in this paper was completed with the support of SPSS software, and model calculation is realized by EXCEL software.

4.2. The Analysis Results of Principal Component. Based on RSEI model, the ecological environment quality model of Dongjiangyuan River watershed was constructed. Firstly, the indexes of humidity, greenness, dryness, and heat were extracted by using remote sensing software, and then the four indexes were normalized and superimposed. The principal component analysis was performed on the superimposed layers, and finally, the ecological environmental quality index of the Dongjiangyuan River source basin was obtained. The method of principal component analysis is adopted to automatically select the characteristic contribution rate of each component, which can effectively avoid the influence of human factors on the evaluation result and make the evaluation result more objective and accurate. Principal component analysis results of indicators in the source basin of the Dongjiangyuan River in each year are shown in Tables 1–4.

As can be seen from the table, the four indicators contributed to PCI, and the contribution rates of characteristic values of PCI were 77.43, 73.65, 78.91, 71.06, 73.47, 77.47, and 75.05, respectively, with all the contribution rates

TABLE 1: Principal component analysis results of indexes in 2004 and 2010.

Index	2004				2010			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
Wet	0.237	-0.022	0.769	0.593	0.227	0.106	0.832	0.495
NDVI	0.632	0.456	-0.486	0.395	0.534	0.567	-0.467	0.420
NDSI	-0.563	-0.225	-0.373	0.702	-0.451	-0.408	-0.287	0.761
LST	-0.476	0.861	0.179	-0.010	-0.701	0.708	0.083	0.029
Characteristic value	0.017	0.003	0.001	0.000	0.011	0.003	0.001	0.000
Eigenvalue contribution rate	78.91	15.17	5.21	0.71	71.06	21.80	6.44	0.70

TABLE 2: Principal component analysis results of indexes in 2013 and 2018.

Index	2013				2018			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
Wet	0.2011	0.031	0.714	0.670	0.183	0.008	0.695	0.696
NDVI	0.607	0.584	-0.460	0.281	0.680	0.492	-0.466	0.282
NDSI	-0.457	-0.252	-0.505	0.687	-0.499	-0.190	-0.528	0.660
LST	-0.618	0.771	0.154	-0.015	-0.506	0.850	0.145	-0.022
Characteristic value	0.012	0.004	0.001	0.000	0.012	0.003	0.001	0.000
Eigenvalue contribution rate	73.47	21.44	4.74	0.35	77.47	17.53	4.64	0.36

TABLE 3: Principal component analysis results of indexes in 1991 and 2000.

Index	1991				2000			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
Wet	0.284	0.070	0.864	0.410	0.184	-0.088	0.762	0.615
NDVI	0.479	0.613	-0.426	0.461	0.509	0.369	-0.536	0.564
NDSI	-0.420	-0.404	-0.203	0.787	-0.713	-0.306	-0.308	0.551
LST	-0.716	0.675	0.176	0.009	-0.446	0.873	0.195	0.017
Characteristic value	0.011	0.002	0.001	0.000	0.013	0.003	0.001	0.000
Eigenvalue contribution rate	77.43	16.66	5.23	0.68	73.65	19.16	6.09	1.10

TABLE 4: Principal component analysis results of indexes in 2019.

Index	2019			
	PC1	PC2	PC3	PC4
Wet	0.160	0.012	0.719	0.676
NDVI	0.663	0.537	-0.432	0.293
NDSI	-0.464	-0.221	-0.528	0.676
LST	-0.565	0.814	0.131	-0.020
characteristic value	0.013	0.004	0.001	0.000
Eigenvalue contribution rate	75.05	20.36	4.31	0.28

greater than 70%, indicating that the first principal component integrated most of the information of the four remote sensing ecological factors. The value of humidity index and greenness index is positive sign, indicating that humidity and greenness have positive effect on ecological environment. The value of dryness index and heat index is negative, indicating that dryness and heat have negative effect on ecological environment.

4.3. Variation Characteristics of RSEI Index in the Headwaters of Dongjiangyuan River

4.3.1. *The Variation Characteristics of Humidity Index.* The statistical humidity index of remote sensing data in each period is shown in Table 5.

The wet value is standardized and ranges between [0,1]. In 1991, 2000, 2004, 2010, 2013, 2018, and 2019, the mean

TABLE 5: Statistical table of humidity index of remote sensing data in each period.

Humidity indicators	Minimum value	Maximum value	Mean value	Standard deviation
1991	0.00	1.00	0.788	0.039
2000	0.00	1.00	0.766	0.035
2004	0.00	1.00	0.791	0.041
2010	0.00	1.00	0.769	0.037
2013	0.00	1.00	0.840	0.031
2018	0.00	1.00	0.810	0.028
2019	0.00	1.00	0.778	0.028

values of wet in Dongjiangyuan River were 0.788, 0.766, 0.791, 0.769, 0.840, 0.810, and 0.778, respectively. The humidity index fluctuated in a small range. But the overall trend is up.

The wet distribution maps of the Dongjiangyuan River in each period are shown in Figure 1.

From the spatial distribution map of humidity, it can be seen that the areas with low humidity are mainly distributed in Wenfeng, Changning, Nanqiao, Liuche, Shuiyuan, Zhengang, Kongtian, Hezi, Longtang, Lishi, Tianjiu, Kui Mei Mountain, old town, and township areas. The area covered by vegetation has a higher value of humidity. The distribution of humidity is closely related to human activities.

4.3.2. The Variation Characteristics of Greenness Index. The statistical table of the greenness index of remote sensing data in the source basin of the Dongjiangyuan River in each period is shown in Table 6.

NDVI values are standardized and range between [0, 1]. As can be seen from the table, the mean NDVI values in 1991, 2000, 2004, 2010, 2013, 2018, and 2019 were 0.851, 0.878, 0.768, 0.855, 0.876, 0.884, and 0.887, respectively. The NDVI value increased from 1991 to 2000 but decreased from 2000 to 2004 due to the destruction of ecological vegetation and increased continuously from 2004 to 2019 in the source basin of the Dongjiangyuan River. The NDVI value in the source basin of Dongjiangyuan River showed a trend of first increasing, then decreasing, and then increasing. Because the government issued a series of policies to protect the ecological environment and took vegetation restoration measures, the NDVI value in the source basin of Dongjiangyuan River is increased.

The distribution diagram of NDVI in the source watershed of the Dongjiangyuan River is shown in Figure 2.

From the analysis of the angle of spatial distribution, Wenfeng, Changning, Nanqiao, Liuche, Shuiyuan, Zhengang, Kongtian, Tachi, Longtang, Lishi, Tianjiu, Kui Mei Mountain, and Laocheng towns show red or yellow, and NDVI is low. The ecological vegetation in these areas has been destroyed by human activities. From 1991 to 2000, the red and yellow areas decreased, indicating that the ecological status of Dongjiangyuan River improved. From 2000 to 2004, the area of red and yellow regions increased, and the ecological environment of some regions deteriorated. From 2004 to 2010, the area of the green zone in Dongjiangyuan River increased and the ecological environment improved.

From 2010 to 2013, the scope of green areas in the Dongjiangyuan River further increased, and the ecological environment continued to improve. In 2013, 2018, and 2019, the NDVI of Dongjiangyuan River showed no significant spatial change. From 1991 to 2019, the ecological vegetation status in the Dongjiangyuan River became better.

4.3.3. The Variation Characteristics of Dryness Index. The statistical table of dryness index for remote sensing data in each period in the source basin of the Dongjiangyuan River is shown in Table 7.

NDSI values are standardized and range between [0, 1]. As it can be seen from the table, the mean NDSI values in 1991, 2000, 2004, 2010, 2013, 2018, and 2019 were 0.477, 0.458, 0.623, 0.470, 0.445, 0.434, and 0.432, respectively. The NDSI value decreased from 1991 to 2000, increased from 2000 to 2004, and decreased from 2004 to 2019 in the Dongjiangyuan River. From 1991 to 2019, NDSI values in the Dongjiangyuan River showed a trend of decreasing, then increasing, and decreasing again. From 1991 to 2000, the main reason for the decrease of NDSI was the increase of vegetation area and the decrease of bare land area. From 2000 to 2004, the ecological environment of Dongjiangyuan River was seriously damaged, and the area of bare land increased sharply, which led to the increase of dryness index. The reason why the dryness index of Dongjiangyuan River decreased continuously from 2004 to 2019 is that the government introduced a series of measures to protect the ecological environment, strengthened the management of rare earth mining, and effectively protected the ecological vegetation.

The NDSI distribution of remote sensing data in the source basin of the Dongjiangyuan River is shown in Figure 3.

In terms of space, Wenfeng, Changning, Nanqiao, Liuche, Shuiyuan, Zhengang, Kongtian, Hezi, Longtang, Lishi, Tianjiu, Kui Mei Mountain, and Laocheng towns show red or yellow, and the large NDSI value indicates a high degree of drying. From 1991 to 2000, the red and yellow areas in the source basin of Dongjiangyuan River decreased, indicating that the degree of drying decreased. From 2000 to 2004, the region of red and yellow regions increased, and the degree of local drying increased. From 2004 to 2010, the area of green area in Dongjiangyuan River is increased, and the degree of drying in the basin is decreased. From 2010 to 2019, the area of the green zone in Dongjiangyuan River kept increasing, and the degree of drying in the basin continued

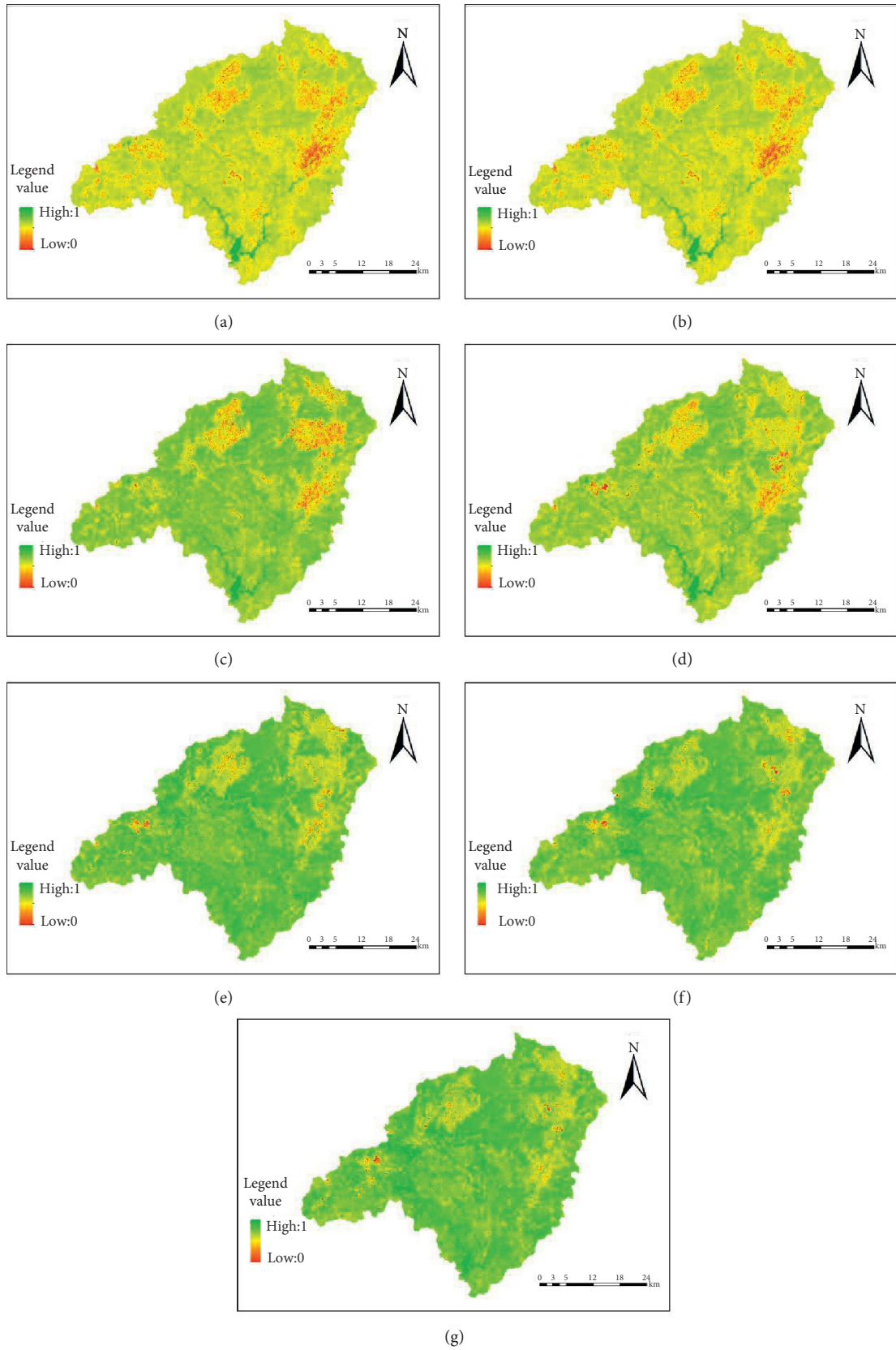


FIGURE 1: The wet distribution map of remote sensing data in various periods for Dongjiangyuan River: (a) 1991; (b) 2000; (c) 2004; (d) 2010; (e) 2013; (f) 2018; (g) 2019.

TABLE 6: Statistical table of greenness index of remote sensing data in each period.

Green degree index	Minimum value	Maximum value	Mean value	Standard deviation
1991	0.00	1.00	0.851	0.061
2000	0.00	1.00	0.878	0.066
2004	0.00	1.00	0.768	0.089
2010	0.00	1.00	0.855	0.068
2013	0.00	1.00	0.876	0.077
2018	0.00	1.00	0.884	0.080
2019	0.00	1.00	0.887	0.084

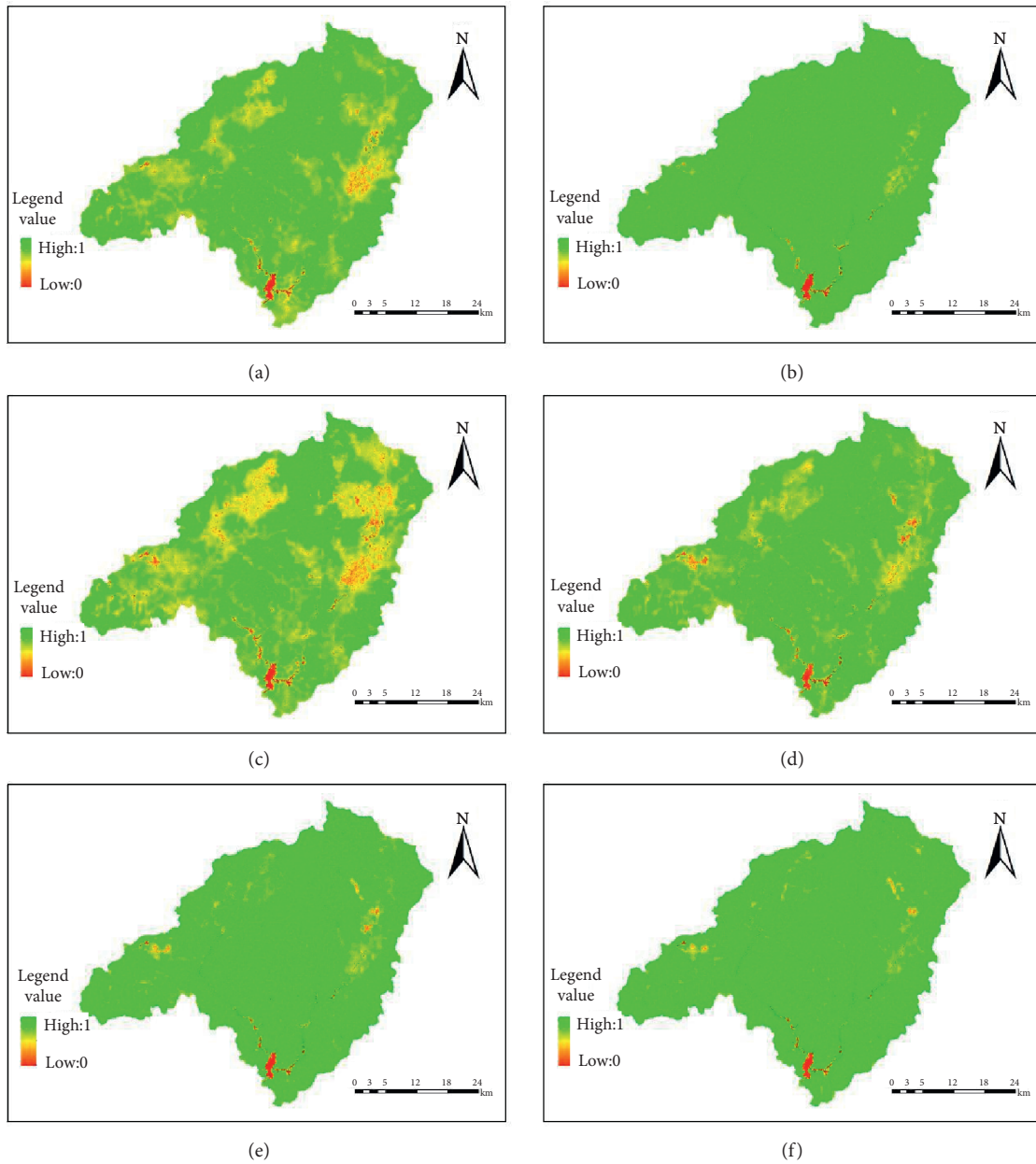
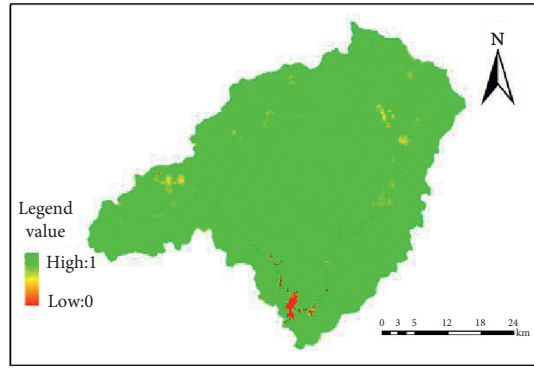


FIGURE 2: Continued.

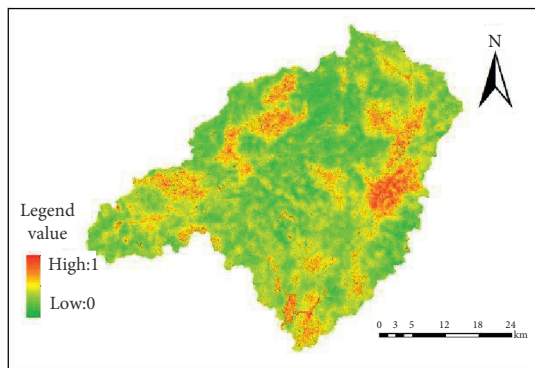


(g)

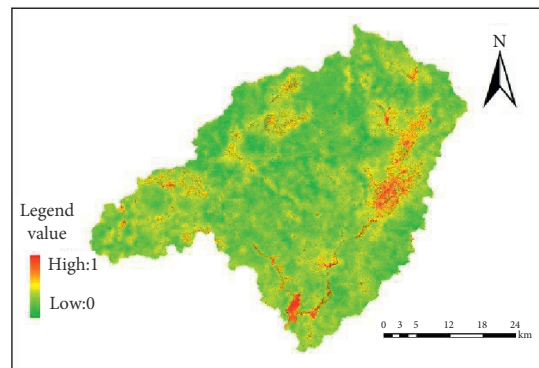
FIGURE 2: The NDVI distribution map of remote sensing data in various periods for Dongjiangyuan River: (a) 1991; (b) 2000; (c) 2004; (d) 2010; (e) 2013; (f) 2018; (g) 2019.

TABLE 7: Statistical table of dryness index of remote sensing data in each period.

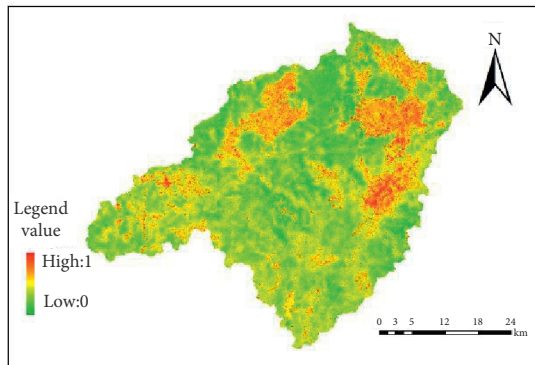
Dryness index	Minimum value	Maximum value	Mean value	Standard deviation
1991	0.00	1.00	0.477	0.050
2000	0.00	1.00	0.458	0.0459
2004	0.00	1.00	0.623	0.077
2010	0.00	1.00	0.470	0.052
2013	0.00	1.00	0.445	0.055
2018	0.00	1.00	0.434	0.058
2019	0.00	1.00	0.432	0.058



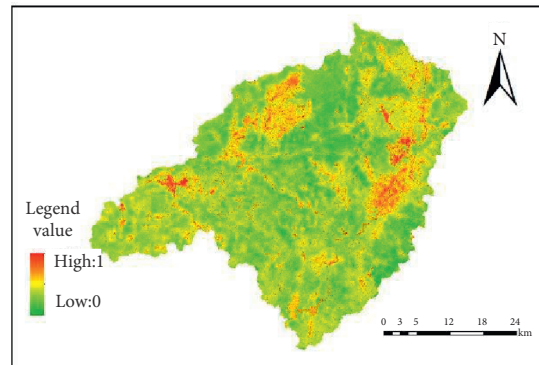
(a)



(b)



(c)



(d)

FIGURE 3: Continued.

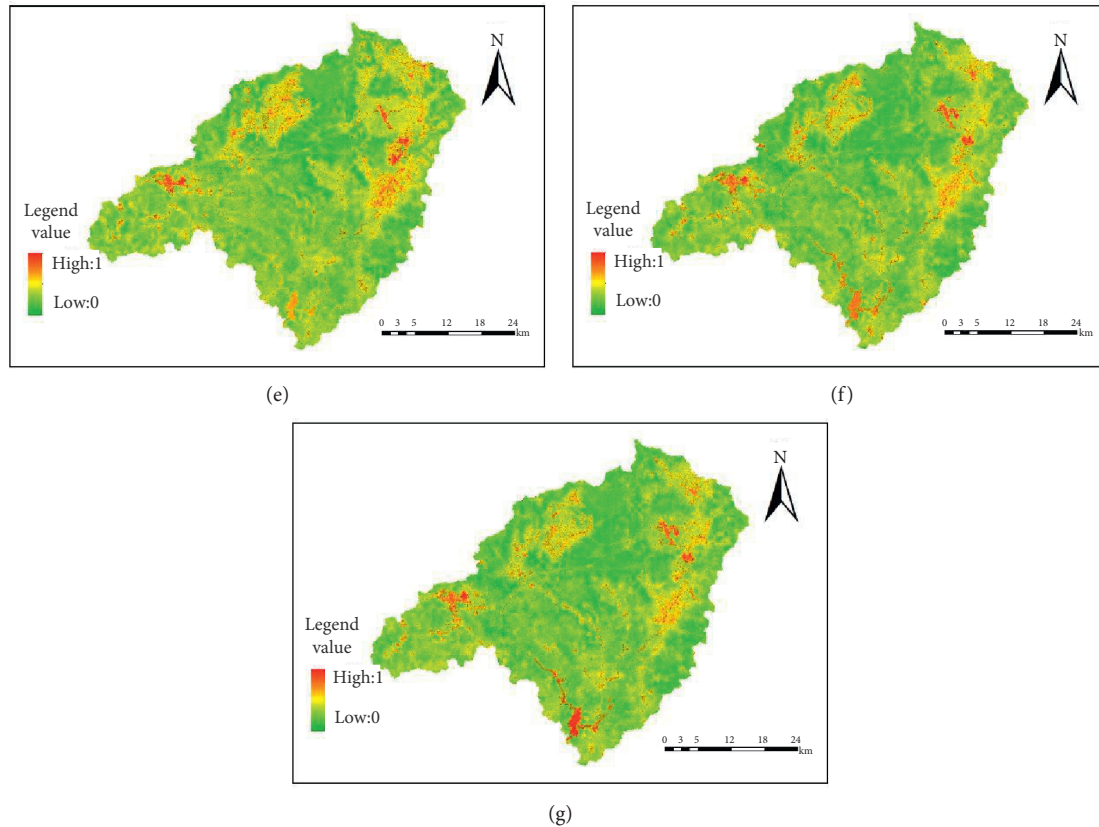


FIGURE 3: The NDSI distribution map of remote sensing data in various periods for Dongjiangyuan River: (a) 1991; (b) 2000; (c) 2004; (d) 2010; (e) 2013; (f) 2018; (g) 2019.

TABLE 8: Statistical table of heat index of remote sensing data in each period.

Heat index	Minimum value	Maximum value	Mean value	Standard deviation
1991	12.322	34.202	20.298	1.781
2000	13.274	37.514	23.247	1.836
2004	14.218	34.430	21.577	1.678
2010	4.909	25.688	14.163	1.949
2013	15.001	38.442	24.922	1.938
2018	12.198	34.556	23.231	1.599
2019	18.372	41.317	26.207	1.883

to be improved. On the whole, the degree of drying in the Dongjiangyuan River continued to be improved from 1991 to 2019.

4.3.4. *The Variation Characteristics of Heat Index.* The statistical table of heat index gained from the remote sensing data in the Dongjiangyuan River in different periods is shown in Table 8.

The mean temperatures in 1991, 2000, 2004, 2010, 2013, 2018, and 2019 were 20.298, 23.247, 21.577, 14.163, 24.922, 23.231, and 26.207, respectively. The temperature has a rising trend as a whole. The average temperature was lower compared with other years in 2010, which is due to the late time of remote sensing images adopted.

The LST distribution map of remote sensing data in various periods of the Dongjiangyuan River is shown in Figure 4.

In terms of space, the red area is mainly distributed in towns and villages around Wenfeng, Changning, Nanqiao, Liuche, Longyan, Chenguang, Changpu, Shuiyuan, Jitan, Zhengang, Kongtian, Keci, Longtang, Lishi, Tianjiu, Kui Mei Mountain, and Laocheng.

4.4. *The Calculation Results of the Ecoenvironmental Quality Index (RSEI).* Based on Landsat remote sensing image processing, the ecoenvironmental quality index was calculated. The distribution of remote sensing ecological index RSEI in 1991, 2000, 2004, 2010, 2013, 2018, and 2019 in

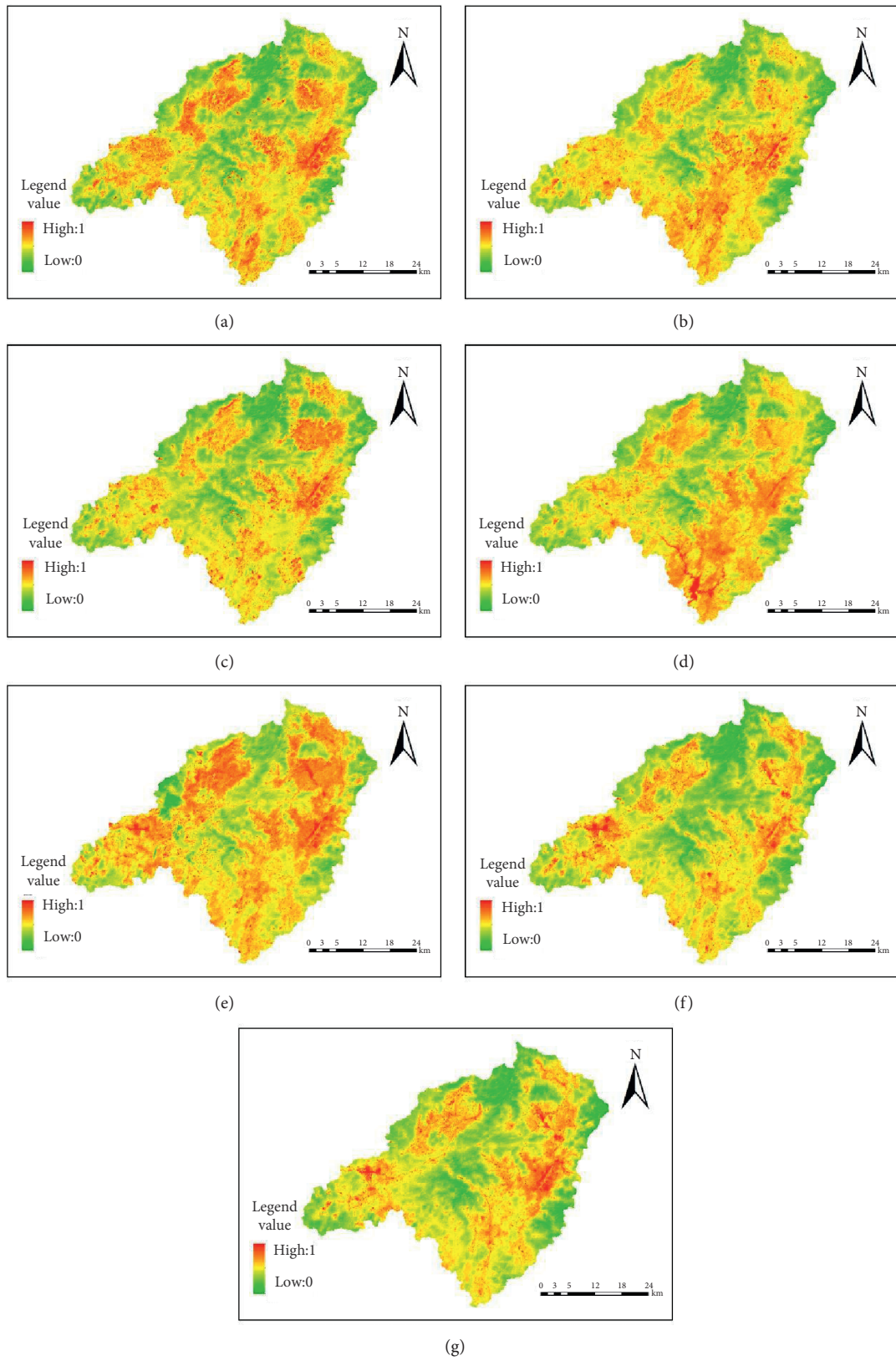


FIGURE 4: The LST distribution map of remote sensing data in various periods for Dongjiangyuan River: (a) 1991; (b) 2000; (c) 2004; (d) 2010; (e) 2013; (f) 2018; (g) 2019.

Dongjiangyuan River was 0.628, 0.653, 0.511, 0.540, 0.689, 0.671, and 0.772, respectively. From 1991 to 2019, the ecoenvironmental quality index of Dongjiangyuan River watershed showed a trend of increasing first, decreasing, and then increasing. From 1991 to 2000, vegetation coverage increased and ecological environment quality improved. From 2000 to 2004, the area of bare land increased greatly and the quality of ecological environment deteriorated. From 2004 to 2019, vegetation coverage increased and ecological environment quality improved. On the whole, the ecological environment of Dongjiangyuan River basin is developing in a good direction.

5. Conclusion

In this study, seven remote sensing images of 1991, 2000, 2004, 2010, 2013, 2018, and 2019 were selected from the source basin of the Dongjiangyuan River, and the indexes of humidity, greenness, dryness, and heat were extracted by remote sensing processing software. The ecological environment quality evaluation model was built by principal component analysis of four indexes, and the mean values of remote sensing ecological indexes in different years were calculated to analyze the changing trend of ecological environment quality in Dongjiangyuan River. The following conclusions were drawn:

- (1) From 1991 to 2019, the ecoenvironmental quality of Dongjiangyuan River showed an uptrend, downturn, and uptrend, and the ecoenvironmental quality was developing in a good direction overall.
- (2) From 1991 to 2019, the humidity index of the Dongjiangyuan River fluctuated in a small range.
- (3) From 1991 to 2019, the greenness index in the source basin of the Dongjiangyuan River showed a trend of increasing first, decreasing then, and increasing again.
- (4) From 1991 to 2019, the dryness index in the Dongjiangyuan River first decreased, then increased, and decreased again; the average temperature in the source basin of the Dongjiangyuan River showed an increasing trend.
- (5) The change of the land role was an important factor that affected ecoenvironmental quality. The mean values of remote sensing ecological indices of different land types were compared, where the RSEI values of forestland were higher than those of other land types, and the change of the role of land will affect the change of ecological environment quality.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] M. Ibrahim, A. El-Zaarta, and C. Adams, "Smart sustainable cities roadmap: readiness for transformation towards urban sustainability," *Sustainable Cities and Society*, vol. 37, pp. 530–540, 2018.
- [2] J. S. Alawamy, S. K. Balasundram, A. H. Hanif, and C. T. Boon Sung, "Detecting and analyzing land use and land cover changes in the region of Al-jabal Al-akhdar, Libya using time-series Landsat data from 1985 to 2017," *Sustainability*, vol. 12, no. 11, p. 4490, 2020.
- [3] S. Chang, Z. Wang, D. Mao, K. Guan, M. Jia, and C. Chen, "Mapping the essential urban land use in changchun by applying random forest and multi-source geospatial data," *Remote Sensing*, vol. 12, no. 15, p. 2488, 2020.
- [4] R. P. Bartelme, S. L. McLellan, and R. J. Newton, "Freshwater recirculating aquaculture system operations drive biofilter bacterial community shifts around a stable nitrifying consortium of ammonia-oxidizing archaea and comammox Nitrospira," *Frontiers in Microbiology*, vol. 8, pp. 1–18, 2017.
- [5] Z. Han, Y. Zhang, W. An, J. Lu, J. Hu, and M. Yang, "Antibiotic resistomes in drinking water sources across a large geographical scale: multiple drivers and co-occurrence with opportunistic bacterial pathogens," *Water Research*, vol. 183, pp. 1–11, 2020.
- [6] K. Biegert, D. Stöckeler, R. J. McCormick, and P. Braun, "Modelling soluble solids content accumulation in "braeburn" a," *Plants*, vol. 10, no. 2, p. 302, 2021.
- [7] C. B. Watkins and J. P. Mattheis, *Apple. Postharvest Physiological Disorders in Fruits and Vegetables*, CRC Press, Boca Raton, FL, USA, 2019.
- [8] L. Liu, X. Huang, A. Ding, and C. Fu, "Dust-induced radiative feedbacks in north China: a dust storm episode modeling study using WRF-Chem," *Atmospheric Environment*, vol. 129, pp. 43–54, 2016.
- [9] D. Zhou, K. Ding, X. Huang et al., "Transport, mixing and feedback of dust, biomass burning and anthropogenic pollutants in eastern Asia: a case study," *Atmospheric Chemistry and Physics*, vol. 18, no. 22, pp. 16345–16361, 2018.
- [10] M. Beltrán-Esteve, E. Reig-Martínez, and V. Estruch-Guitart, "Assessing eco-efficiency: a metafrontier directional distance function approach using life cycle analysis," *Environmental Impact Assessment Review*, vol. 63, pp. 116–127, 2017.
- [11] P. Peters, "Mobilizing rural low-income communities to assess and improve the ecological environment," *Journal of Nutrition Education and Behavior*, vol. 45, no. 4, pp. 443–451, 2013.
- [12] A. Tuya, J. Ren, and B. Yin, "Research on the classification of coupling relationship among the ecological environment, natural disasters and regional poverty in inner Mongolia autonomous region," *Applied Mechanics and Materials*, vol. 260–261, pp. 1063–1068, 2013.
- [13] R. O. Yin, B. Heidrich, S. Oyewole, O. T. Okareh, and C. W. McGlothlin, "Chemo metric analysis of ecological toxicants in petrochemical and industrial environments," *Chemosphere*, no. 36, pp. 112–119, 2014.
- [14] V. Remig, "Mobilizing rural low-income communities to assess and improve the ecological environment to prevent childhood obesity," *Journal of Nutrition Education and Behavior*, vol. 48, no. 7, pp. 1135–1141, 2016.
- [15] M. Ashfaq and K. Nawaz, "Ecological risk assessment of pharmaceuticals in the receiving environment of pharmaceutical wastewater in Pakistan," *Ecotoxicology and Environmental Safety*, vol. 136, 2017.

- [16] D. Hou, D. O'Connor, and A. D. Igalavithana, "Metal contamination and bioremediation of agricultural soils for food safety and sustainability," *Nature Reviews Earth & Environment*, vol. 1, pp. 366–381, 2020.
- [17] Y. H. Zhang, J. J. Bi, K. X. Gao, and Y. J. Hou, "Study on ecological environment assessment and dynamic monitoring of Tai'an city based on remote sensing and GIS," *Journal of Shandong Meteorology*, vol. 32, no. 1, pp. 1–5, 2012.
- [18] H. Q. Xu, "A remote sensing index for assessment of regional ecological changes," *China Environmental Science*, vol. 33, no. 5, pp. 889–897, 2013.
- [19] F. M. Jia, S. C. Guo, J. F. Bai, and B. C. Zhang, "Dynamic assessment of ecological environment quality," *Hunan Agricultural Sciences*, vol. 4, no. 6, pp. 60–62, 2014.
- [20] X. X. Huang, B. Z. Chen, and N. Liu, "Ecological environment analysis of mining area based on fuzzy comprehensive evaluation model," *Nonferrous Metals*, vol. 63, no. 6, pp. 68–71, 2015.
- [21] H. Liu, G. H. Che, C. J. Gao, and Y. J. Zhao, "Ecological environment quality change and its influencing factors in Hebei Province," *Ecological Science*, vol. 35, no. 2, pp. 89–97, 2016.
- [22] Y. F. Gao, X. Chen, and Z. H. Xin, "Comprehensive evaluation of Ecological environment quality in Xi'an city," *Statistics and Management*, no. 7, pp. 55–57, 2017.
- [23] X. X. Li and H. Luo, "Ecological environment quality assessment of Beipanjiang basin based on GIS and remote sense," *Environmental Science and Management*, vol. 43, no. 7, pp. 178–182, 2018.
- [24] H. Z. Li, "Temporal and spatial changes of ecological environment quality in Chaohu Lake Basin based on RSEI," *Journal of Anhui Agricultural Sciences*, vol. 47, no. 16, pp. 82–86, 2019.
- [25] C. Xu, L. Cao, and Y. Yuan, "Research on ecological environment quality of Anhui province based on principal component analysis," *Resources Development and Market*, vol. 24, no. 2, pp. 118–119+179, 2008.
- [26] Y. Q. Cheng and G. T. Chi, "Ecological evaluation model based on kernel principal component analysis and its application research," *Chinese Management Science*, vol. 19, no. 3, pp. 182–192, 2011.
- [27] D. S. Sun, "Statistical analysis of environmental quality in nanjing based on principal component analysis," *Resource conservation and environmental protection*, vol. 19, no. 3, pp. 182–192, 2011.
- [28] W. J. Li and Q. D. Tian, "Comprehensive evaluation of eco-city construction based on principal component analysis: a case study of xi'an city, shaanxi province," *Economic Perspectives*, vol. 1, no. 1, pp. 11–14, 2015.
- [29] B. G. Wang, "Study on environmental quality evaluation index system of county (city) region," *Construction of Small towns*, vol. 1, no. 1, pp. 75–79, 2016.

Research Article

Deep Learning-Based Fake Information Detection and Influence Evaluation

Ning Xiang 

School of Journalism and Communication, Hunan Mass Media Vocational and Technical College, Changsha 410100, China

Correspondence should be addressed to Ning Xiang; 9120130138@jxust.edu.cn

Received 10 January 2022; Accepted 26 January 2022; Published 23 February 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Ning Xiang. 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 prevalence of the Internet, a large number of users have participated in OSN (Online Social Networks), which has gradually made it the mainstream way for obtaining news or information from the Internet. However, with the rapid development of the Internet, a large amount of fake information has also been spread on the Internet. Therefore, fake information detection is of great significance at the moment. A multimodal fake information detection method is proposed in this article, which has adopted the textual and visual contents in the piece of information to make the judgments. The textual feature representation vector is firstly obtained through the pretraining of the Bert model, and then the visual feature representation is obtained through the pretraining of the VGG-19 model. From the proposed method, two MCBP (Multimodal Compact Bilinear Pooling) modules are adopted. The first MCBP module is adopted to obtain the visual feature representation vector with attention, and the second MCBP module is adopted to join the visual feature with the attention mechanism and the textual feature vector. Then, the joined vector can be adopted for fake information detection. The proposed method in this article is compared with two baseline methods. The experimental results on the Twitter and Weibo datasets have proved that the proposed method in this article is better than the EANN method and the SpotFake method in terms of accuracy, precision, recall, and F1 score.

1. Introduction

With the popularity of the Internet, a large number of users have participated in OSN (Online Social Networks), which has gradually made it the mainstream method for obtaining news or information from the Internet. However, with the rapid development of the Internet, much fake information was also spread on the Internet. In particular, the creation of the programmed social accounts, called Socialbots [1–3], has flooded the OSN with information generated from Socialbots. According to a report by the company GlobalDots, the traffic generated by Socialbots on social platforms in 2018 has reached about 37% of the total traffic, of which malicious traffic accounted for about 20% [4]. From the overall traffic, we can see the severity of fake information. On the other hand, the propagation of some fake information has seriously impacted society, which can even affect political elections and manipulate global stock markets. For example, in 2016, the fake news storm has affected the US election. In

2017, a large amount of spam information was generated by Socialbots during the election of the German Chancellor. Also, in 2017, a large amount of fake information was generated by Socialbots just before the French presidential election [5–7]. In addition, with the spread of the new COVID epidemic, all kinds of fake information are also flooded on the Internet, which has a serious impact on the lives of the public. Therefore, fake information detection is an important topic at present, and blocking the spread of fake information is of great significance to the normal operation of society.

The methods of fake information detection can be described mainly from two aspects, the first is the relevant features for fake information detection, and the second denotes the model. Overall, the features for fake information detection can be divided into two types: one denotes the features from the content and the other denotes the contextual feature. In the method of content-based fake information detection, it is performed by extracting relevant

features in fake information, such as textual features and visual features [8–10]. Textual features generally consisted of word frequency, special vocabulary, total number of words, and syntactic features. Visual features often include sharpness, histogram features, and consistency features. The difference between the contextual features and the content features is that the contextual features are not limited to the current piece of information. In fact, the contextual features can describe the information within the propagation of the information. Normally, it includes the following aspects. On the user characteristic level, it includes the information forwarded by the user, including the user’s attention and the user’s registration time. The other is within the process when the user forwards the information, including the stance network, friend network, forwarding time, and forwarding group. As the fake information detection methods based on contextual features involve the propagation model and require a larger amount of data, it will not be studied in this article. The fake information detection methods herein only include content-based detection methods; that is, only the content for this piece of information is adopted for detection.

The features for content-based fake information detection can usually be divided into textual features and visual features. The extraction of textual features and visual features are introduced as follows. The representation methods of textual features include the bag-of-words method [11, 12], which treats the text as a collection of independent vocabulary, ignoring the grammatical and word sequences in the text. For the TF-IDF method [13, 14], the importance of different vocabulary to the overall text is considered. It is assumed that the importance of the vocabulary is proportional to the frequency of appearance in the text and inversely proportional to the frequency of appearance in the corpus. In the Word2Vec method [15, 16], the low-dimensional word embedding can be learned through a shallow neural network according to the corpus. The Word2Vec method includes the CBOW model and the skip-gram model. In the skip-gram model, the importance of words that are closer is greater than those farther away, which is different from the CBOW model. In practical application, the Word2Vec method is usually adopted as a pretraining method to generate the corresponding word embedding vector. In this representation, the underlying information in the existing corpus can be fully utilized. In this condition, words with similar semantics also have a closer metric in the word embedding vector space. The Bert model [17, 18] is transformer-based and is a deep, bidirectional, and unsupervised representation model. This model can sufficiently adapt contextual information. After training through a large number of the corpus, ambiguities can be weakened compared to methods such as Word2Vec. At present, the Bert model has the best representation capability. For visual vector representation, it is currently relatively mature, including ResNet [19], VGGNet [20] with repeated modules, GoogleLeNet [21], which is a fusion of CCN and RNN models, and AlexNet [22]. These models are relatively mature, so they are not repeated here.

Content-based fake information detection methods can be divided into single-modal-based methods and multimodal methods according to the types of content adopted. Among the single-modal-based methods, only a single source of information, such as text information or visual information, is adopted for detection. In the multimodal-based methods, multiple types of information are joined for comprehensive detection. For example, texts and images are adopted for detection. Generally speaking, multimodal-based detection methods have higher accuracy than single-modal-based. For the single-modal-based methods, the authors in [23] proposed to adopt the CNN (Convolutional Neural Network) to learn the deep features in the information to identify true or false. The authors in [24] proposed to detect the authenticity of information through a RNN (Recurrent Neural Network). In [25], by adopting the rich information from the image, fake information detection is performed through a multiregion neural network. In single-modal-based methods, since just a single type of information is adopted, the accuracy can be improved. In multimodal-based methods, textual representation vectors and visual representation vectors can be concatenated to obtain a hybrid representation vector for fake information detection. In the method proposed by Jin et al. [26], an attention-based RNN model is adopted to fuse textual features and visual features, and good detection results can be achieved. In literature [27], an adversarial network is adopted to obtain a multimodal feature extractor so as to eliminate the influence of common features in different information on detection, which can improve the generalization ability of the proposed model. This method can be called the EANN method, where direct concatenation is still applied to fuse the visual and textual feature vectors. The SpotFake model is proposed in literature [28], where the textual feature vector is extracted through a pretrained Bert model, and the visual feature vector is extracted through the pretrained VGG-19 model. Then the two vectors are concatenated to obtain a new feature vector for detection. The SpotFake+ model is proposed in literature [29]. Compared with the SpotFake model, it has adopted the improved Bert model XLNet to extract visual feature vectors, which can improve the detection performance compared to the SpotFake method. The authors in [30] have proposed the SAME method, where sentiment classification results are adopted for fake information detection.

A multimodal-based fake information detection method is proposed in this article, which adopts the textual information and visual information from the piece of information to make judgments. Firstly, the textual feature representation vector is obtained through the pretraining of the Bert model, and then the visual feature representation of the image can be obtained through the pretraining of the VGG-19 model. In order to better fuse the two different types of features, instead of simple concatenation, Multimodal Compact Bilinear Pooling (MCBP) is adopted for fusing features in the two different domains. The proposed method in this article is inspired by the Visual Question Answering (VQA) method in literature [31]. The MCBP module is used twice here. The first MCBP module is adopted to obtain the

image feature representation with attention, and the second MCBP module is adopted to join the image feature with the attention mechanism and the textual feature vector. And then, fake information detection is performed adopting the joined vectors. The MCBP module adopted in this article can fully integrate the feature vectors of different domains, thereby reducing information loss, at the same time achieving a better computational burden and better back-propagation characteristics. Two OSN datasets, including the Twitter dataset and the Weibo dataset, are adopted to verify the proposed method. Compared with the direct concatenation-based SpotFake method and the EANN method, the proposed method can achieve better accuracy, precision, recall, and F1 score, which has proved the effectiveness of the proposed method.

2. Methods

The proposed method in this article is shown in Figure 1. Firstly, the representation vector of the text information is obtained through pretraining from adopting the Bert model. The representation vector is used as the inputs of two different subnetworks. The first subnetwork is an image feature extraction network based on attention, and the second subnetwork is a feature fusion network of different domains. In the attention-based image feature extraction subnetwork, the MCBP module is adopted for the first time to fuse the visual features and the textual features to obtain the attention vector, which can reweight the visual features directly extracted through VGG-19, thereby obtaining the visual feature vector with attention. This attention-based feature and the representation vector of the textual domain based on Bert are then fused in the feature fusion subnetwork through a second MCBP module. The fused feature vector can then pass through the fake information detection subnetwork. Note that the MCBP module is used twice in the proposed method. Both MCBP modules are adopted for feature fusion of the representation vectors from different domains. The first MCBP is adopted to obtain the image representation vector with attention. The second MCBP is to get the feature vector after fusion. The length of the visual and textual representation vector can be regarded as hyperparameters. The hyperparameters are studied in the experiment section. Herein, the dimension of the textual and visual vector is 1024 and the flattened vector after the second MCBP module has a length of 8000. The following describes the proposed method from the following four aspects.

2.1. Textual Feature Extraction. Among the many textual feature extraction methods, the most typical one denotes the bag-of-words method, where the texts are treated as a collection of independent words, ignoring the grammatical and sequence features in the text. In the TF-IDF method, for a word in a specific document w_i , the degree of importance can be expressed as follows:

$$tf_{ij} = \frac{n_{ij}}{\sum_k n_{kj}}, \quad (1)$$

where n_{ij} denotes the number of times the word appears in the document d_j and k represents the total number of words in the document d_j . In this method, the importance of words is considered to be proportional to the frequency of appearance in the text and inversely proportional to the frequency of appearance in the corpus. The classic Word2Vec method, as shown in Figure 2, includes the classic CBOW model and the skip-gram model. Noting that the figure herein is just an illustration figure, which shows the most significant differences between the CBOW and the skip-gram model. For the CBOW model, the contextual words are adopted to predict the central word, while for the skip-gram model, the central word is adopted to predict the contextual word.

In this article, the Bert model is adopted to generate word embedding vectors. As shown in Figure 3, the Bert model in this article is similar to that in [28], with several transformer blocks in the middle, using the words of the moving window as input. The attention mechanism is added to each layer's transformer blocks, and then the attention vector is passed to the transformer block of the next layer.

After the word embedding vectors are obtained by the Bert model, the sequence of the word embedding vectors is passed through a two-layer LSTM layer, where the representation vector of the text sequence can then be obtained.

2.2. VGG-19 Model-Based Visual Feature Extraction. In this article, the pretrained VGG-19 neural network by ImageNet dataset is adopted to extract the visual feature vector. Specifically, similar to literature [28], using the output of the last two layers of the classic VGG-19 neural network, and by adding a fully connected layer, we can get the following:

$$V_{vec} = \sigma(W_{dense}V_{VGG-2}). \quad (2)$$

Among them, W_{dense} represents the weight of the fully connected layer, and V_{vec} represents the output of the fully connected layer, which is adopted as a feature vector extracted from the image for later processing.

2.3. The MCBP Module. When there exists a multimodal feature vector fusion, especially when the feature vectors are extracted from different domains, only element-wise product or sum of two feature vectors, or by direct concatenation, will often lead to the deterioration of the final recognition rate due to loss of relevant information of the features in different domains. In the fusion of multimodal-based textual and visual feature fusion, bilinear pooling is a rather reasonable choice. The two feature representations can be fused through the outer product, so that the fused feature representation is translationally invariant. Assuming that two features A and B can be represented as $f_A(l, I) \in R^{L \times M}$ and $f_B(l, I) \in R^{L \times N}$, respectively, where l represents the position in the image, and $L \times M$ and $L \times N$, respectively, represent the dimension of the feature, then bilinear pooling can be performed as follows:

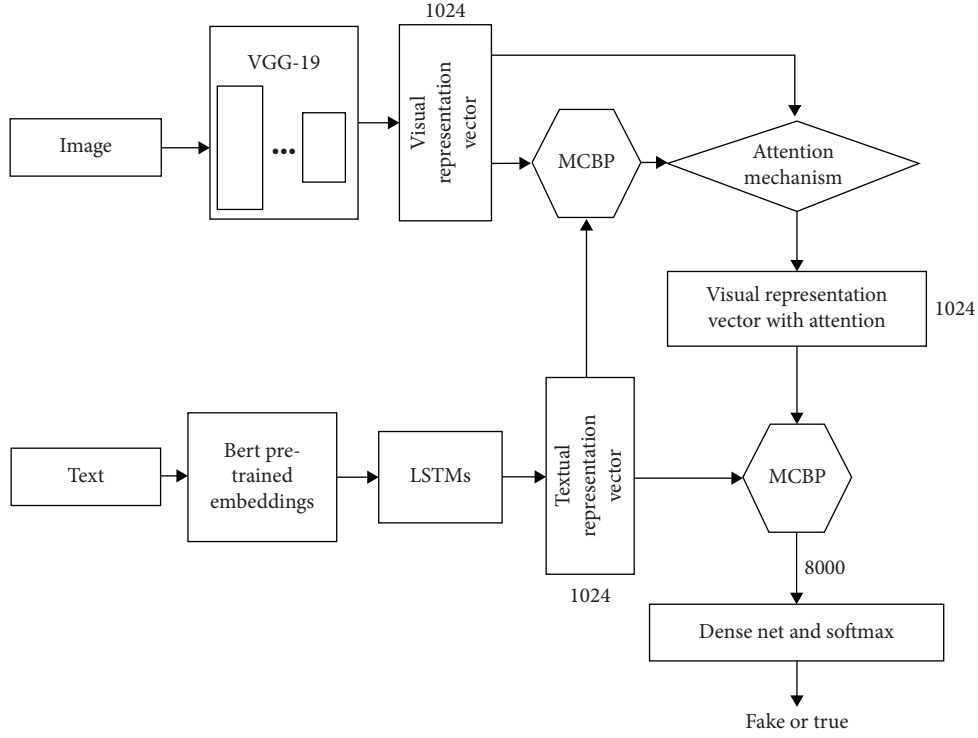


FIGURE 1: The overall structure of the proposed fake information detection method.

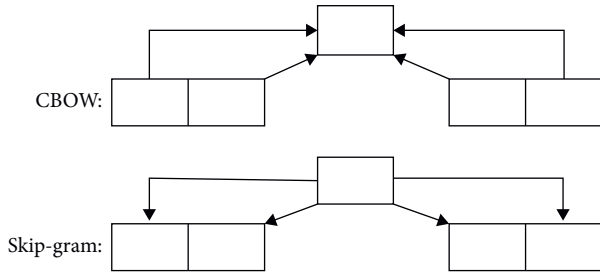


FIGURE 2: Schematic diagram of CBOV model and skip-gram model.

$$\begin{aligned}
 b(l, I, f_A, f_B) &= f_A(l, I)^T f_B(l, I)^T \\
 \zeta(I) &= \sum_l b(l, I, f_A, f_B), \\
 x &= \text{flatten}(\zeta(I)), \\
 y &= \text{sign}(x) \sqrt{|x|}, \\
 y &= \frac{y}{\|y\|_2}.
 \end{aligned} \tag{3}$$

The function $\text{flatten}(\cdot)$ means transforming the matrix into a vector, which is denoted as x . The flattened vector is then normalized to obtain the fused feature y . In this article, due to the reason that the related feature dimensions are relatively high and that the bilinear pooling are used twice, the overall computational burden is relatively large.

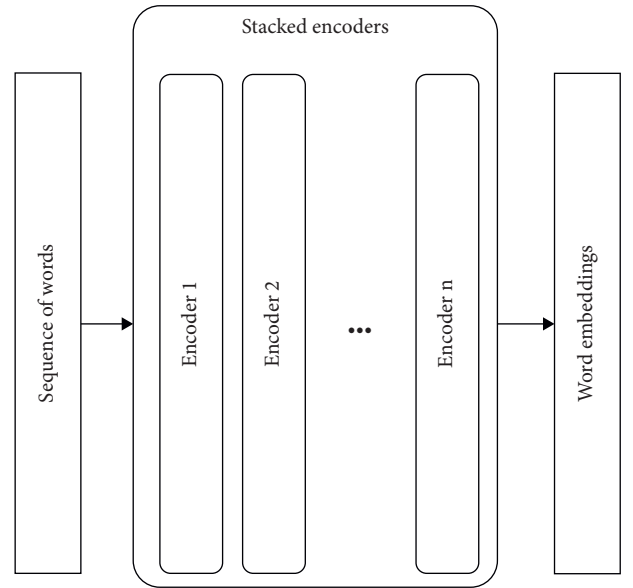


FIGURE 3: An illustration of the Bert-based textual feature extractor.

Therefore, direct use of bilinear pooling is not suitable for the application of this article.

In this article, a new method for fusing two features from different domains is adopted, which has adopted the MCBP module [31]. The basic process of the proposed method is shown in Figure 4. Supposing that the visual feature vector is expressed as $V_a \in R^n$ and the textual feature vector is expressed as $V_b \in R^m$, the structure is shown as follows.

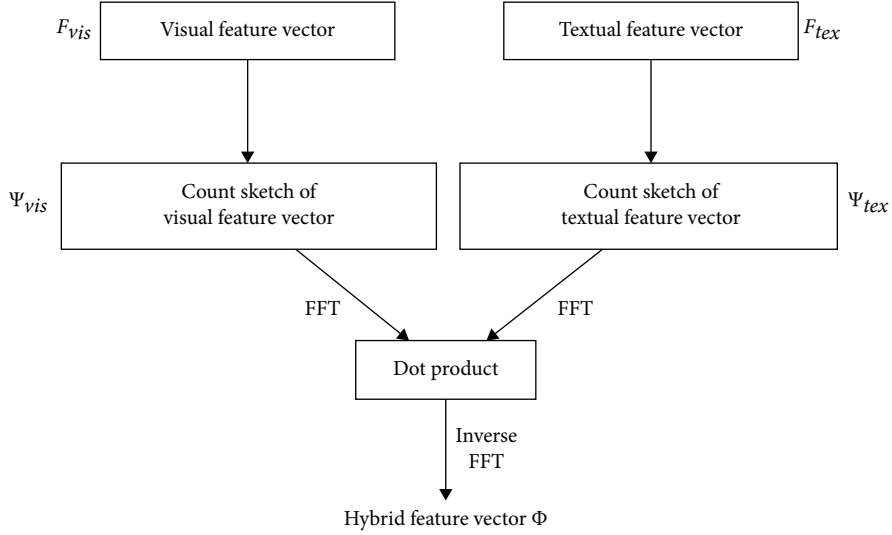


FIGURE 4: The basic process of the MCBP module.

Compared with the previous direct bilinear pooling method, this module has projected the outer product into a low-dimensional space through the count sketch projection function. The projection process adopting the count sketch projection function is shown as follows, which has projected the vector V_a to V_b . Firstly, two vectors are initialized, where

$$\begin{aligned} s &\in \{1, -1\}, \\ h &\in \{1, 2, 3, \dots, m\}^n. \end{aligned} \quad (4)$$

It means that the element in s is either 1 or -1, and h means that the index i in the initial vector V_a is projected to the index j in the vector V_b . Among them, s and h are randomly selected from the uniform distribution. The update process is shown as follows, where the index i is from 1 to n :

$$\begin{aligned} V_a[i]j &= h[i]s[i]V_a[i]V_b[j], \\ V_b[h[i]] &= V_b[h[i]] + s[i]V_a[i]. \end{aligned} \quad (5)$$

Suppose that V_b is initially set as a zero vector. Then according to the above updating formula, in the resulting projecting vector V_b , its corresponding destination index can be expressed as $j = h[i]$, and then $s[i]V_a[i]$ is added to $V_b[j]$.

After the dimensionality reduction projection, the two vectors F_{vis} and F_{tex} are subjected to a convolution operation. In order to speed up the convolution operation, it can be converted into a combination of forward and inverse FFT transformations:

$$\Phi = FFT^{-1}(FFT(F_{vis}) \cdot FFT(F_{tex})). \quad (6)$$

Among them, the operation of \cdot represents element-wise multiplication.

3. Experimental Results

In order to verify the effectiveness of the proposed method in this article on fake information detection, two different datasets are adopted herein. Both datasets are from the OSN

platform: one is the Twitter dataset [32], and the other is the Weibo dataset [26]. Among them, the Twitter dataset consists of a collection of tweets from 17 events and about 2000 other tweets. In the Weibo dataset, all information is manually labeled as true and fake, with the fake information deemed true and already spread on the network. In both the two datasets, the true and fake information is marked, and the detailed statistics of the two datasets are listed in Table 1. In this article, all data samples are divided into training set and test set with a 70%–30% split.

3.1. The Effects of Different Hyperparameters. Two MCBP modules are adopted in the proposed method. The first MCBP module has added an attention mechanism to the visual feature vector, and the second MCBP module integrates visual features and textual features. Table 2 gives the accuracy, precision, recall, and F1 score comparisons under different datasets and under different network settings. Note that here the two different network settings herein are as follows: (1) the proposed structure with two MCBP modules; (2) the single-MCBP structure, where the MCBP module is only adopted to integrate the features from different domains, while there is no attention mechanism added. The structure of the single-MCBP based method is shown as follows. Compared with the proposed double-MCBP based method, the attention-based mechanism is removed and the fused textual and visual vectors are directly adopted for the detection. The structure of the single-MCBP based method is shown in Figure 5.

It can be seen from the table that the proposed double-MCBP module method is better than the single-MCBP method for the two datasets in terms of accuracy, precision, recall, and F1 score. Thus, the proposed double-MCBP module method shows better performance than the single-MCBP module method.

In addition, in the proposed method, the number of feature dimensions adopted for feature fusion also has an important influence on the performance. Here, the number

TABLE 1: The details for the Twitter dataset and the Weibo dataset.

	The Twitter dataset	The Weibo dataset
Fake	7898	4749
True	6026	4779
Total	13924	9528

TABLE 2: The accuracy, precision, recall, and F1 score comparisons under different datasets and under different network settings.

Dataset	Structure	Accuracy	Precision	Recall	F1 score
Twitter	Double-MCBP	0.682	0.596	0.830	0.693
	Single-MCBP	0.551	0.487	0.691	0.571
Weibo	Double-MCBP	0.812	0.833	0.781	0.806
	Single-MCBP	0.633	0.633	0.642	0.637

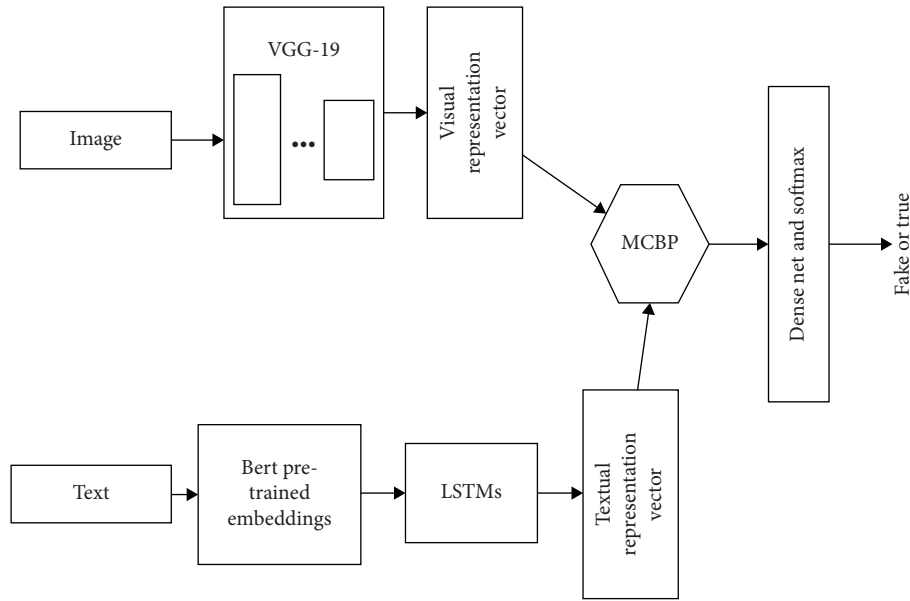


FIGURE 5: The structure of the single-MCBP based method.

of feature dimensions adopted for fusion is denoted as M . Figure 6 shows the changes of accuracies when M takes different values for the Weibo and Twitter datasets. The influences of different values of M are similar to the Weibo dataset. For the Twitter dataset, it can be seen that as the value of M increases from 128 to 1024, the accuracy has increased significantly. Compared with the situations when M equals 128, the accuracy has increased by 14.0%. When M equals 2048, compared with the situation when M equals 1024, accuracy has only increased less than 1%, which is very insignificant. For the Weibo dataset, it can be seen that as the value of M increases from 128 to 1024, the accuracy has increased significantly. Compared with the situations when M equals 128, the accuracy has increased by 19.8%. When M equals 2048, compared with the situation when M equals 1024, accuracy has only increased by about 1%, which is very insignificant. For the two conditions when $M = 2048$ and $M = 1024$, as the parameter size has expanded more than twice, the computational cost needed is also more than twice. Therefore, under the condition that for both the Weibo and Twitter datasets, the increased accuracy is less than 1%, the parameter M is valued as 1024.

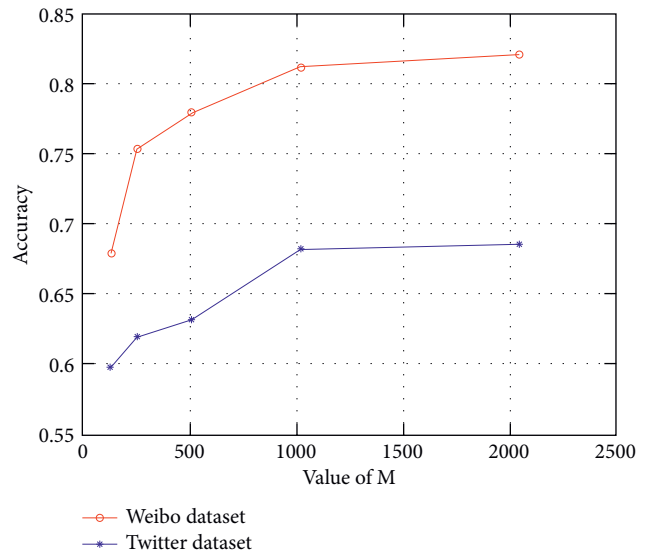
FIGURE 6: Different accuracies with different values of M for the Weibo and Twitter datasets.

TABLE 3: The performance comparisons of the proposed method, the EANN method, and the SpotFake method.

Dataset	Method	Accuracy	Precision	Recall	F1 score
Twitter	EANN	0.623	0.547	0.747	0.631
	SpotFake	0.627	0.550	0.767	0.641
	Proposed	0.682	0.596	0.830	0.693
Weibo	EANN	0.727	0.736	0.711	0.723
	SpotFake	0.769	0.779	0.753	0.766
	Proposed	0.812	0.833	0.781	0.806

3.2. *Comparisons of Different Methods.* In order to illustrate the effectiveness of the proposed method, comparisons with two other methods are made, which can be seen in Table 3. The comparison methods are the EANN method and the SpotFake method. The two methods are selected herein due to the following factors: (1) the proposed method, the EANN method, and the SpotFake method are all multimodal-based methods, which have adopted both textual and visual information for fake information detection; (2) the EANN method and the SpotFake method have adopted direct vector catenation for feature fusion, while the proposed method has adopted the MCBP module to improve feature fusion. The comparison results of the three methods under the Twitter dataset and the Weibo dataset are shown in the following table.

It can be seen that the method proposed is better than the EANN method and the SpotFake method in terms of accuracy, precision, recall, and F1 score. This can fully illustrate that the proposed method in this article can better integrate the features from different domains and has a better performance in fake information detection than methods adopting direct catenation of features from different domains.

4. Conclusions

With the widespread participation of users in OSN, fake information has also been widely spread on the social network and has had a deep impact on social development. Therefore, fake information detection was of great significance at the time when OSN is popularized. A fake information detection method adopting the MCBP module is proposed in this article. In the proposed method, the pre-trained Bert model is adopted to obtain the textual representation vector, and the VGG-19 model is adopted to obtain the visual representation vector. The mentioned MCBP module is adopted twice. For the first time, the MCBP module is used to fuse textual representation vector and visual representation vector to obtain the visual representation vector with attention. Then, the visual representation vector with attention is fused with the textual representation vector to obtain a hybrid representation vector for the classification of true and fake information. The proposed method in this article is compared with the two baseline methods. The results on the Twitter and Weibo open-source datasets prove that the proposed method in this article is better than the EANN method and the SpotFake method in terms of accuracy, precision, recall, and F1 score.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by Education Science “13th five year” Plan project Foundation of Hunan Province, China, “Research on the influence evaluation model of WeChat official account in Higher Vocational Colleges in the era of Convergence Media” (Grant No. XJK19QZY002) and 2017 Project of Think Tank of Hunan Federation of Social Sciences “Research on development countermeasures of Hunan online education industry” (Project No. ZK2017005).

References

- [1] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, “The rise of social bots,” *Communications of the ACM*, vol. 59, no. 7, pp. 96–104, 2016.
- [2] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, “Botornot: a system to evaluate social bots,” in *Proceedings of the 25th international conference companion on world wide web*, pp. 273–274, Montréal, Canada, May 2016.
- [3] Y. Ji, Y. He, X. Jiang, J. Cao, and Q. Li, “Combating the evasion mechanisms of social bots,” *Computers & Security*, vol. 58, pp. 230–249, 2016.
- [4] D. Global, *Bot Traffic Report 2019*, Imperva Incapsula, San Meato, CA, USA, 2019.
- [5] S. C. Woolley, *Automating Power: Social Bot Interference in Global Politics*, First Monday, Canton, TX, USA, 2016.
- [6] A. Bessi and E. Ferrara, “Social bots distort the 2016 US Presidential election online discussion,” *First Monday*, vol. 21, no. 11-7, 2016.
- [7] L. Neudert, B. Kollanyi, and P. N. Howard, *Junk News and Bots during the German Parliamentary Election: What Are German Voters Sharing over Twitter?*, Oxford University, Oxford, UK, 2017.
- [8] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake news detection on social media,” *ACM SIGKDD explorations newsletter*, vol. 19, no. 1, pp. 22–36, 2017.
- [9] E. Tacchini, G. Ballarin, M. L. Della Vedova, S. Moret, and L. de Alfaro, “Some like it hoax: automated fake news detection in social networks,” 2017, <https://arxiv.org/abs/1704.07506>.
- [10] N. Ruchansky, S. Seo, and Y. Liu, “Csi: a hybrid deep model for fake news detection,” in *Proceedings of the 2017 ACM Conference on Information and Knowledge Management*, pp. 797–806, Singapore, November 2017.
- [11] Y. Zhang, R. Jin, and Z. H. Zhou, “Understanding bag-of-words model: a statistical framework,” *International Journal of Machine Learning and Cybernetics*, vol. 1, no. 1-4, pp. 43–52, 2010.
- [12] H. M. Wallach, “Topic modeling: beyond bag-of-words,” in *Proceedings of the 23rd international conference on Machine learning*, pp. 977–984, Pittsburgh, PA, USA, June 2006.

- [13] A. Aizawa, "An information-theoretic perspective of tf-idf measures," *Information Processing & Management*, vol. 39, no. 1, pp. 45–65, 2003.
- [14] W. Zhang, T. Yoshida, and X. Tang, "A comparative study of TF*IDF, LSI and multi-words for text classification," *Expert Systems with Applications*, vol. 38, no. 3, pp. 2758–2765, 2011.
- [15] K. W. Church, "Word2Vec," *Natural Language Engineering*, vol. 23, no. 1, pp. 155–162, 2017.
- [16] X. Rong, "word2vec parameter learning explained," 2014, <https://arxiv.org/abs/1411.2738>.
- [17] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "Bert: pre-training of deep bidirectional transformers for language understanding," 2018, <https://arxiv.org/abs/1810.04805>.
- [18] Y. Liu, M. Ott, N. Goyal et al., "Roberta: a robustly optimized bert pretraining approach," 2019, <https://arxiv.org/abs/1907.11692>.
- [19] 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.
- [20] L. Wang, S. Guo, W. Huang, and Y. Qiao, "Places205-vggnet models for scene recognition," 2015, <https://arxiv.org/abs/1508.01667>.
- [21] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, Boston, MA, USA, June 2015.
- [22] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size," 2016, <https://arxiv.org/abs/1602.07360>.
- [23] F. Yu, Q. Liu, S. Wu, L. Wang, and T. Tan, "A convolutional approach for misinformation identification," in *Proceedings of the IJCAI*, pp. 3901–3907, Melbourne, Australia, August 2017.
- [24] J. Ma, W. Gao, Z. Wei, Y. Lu, and K. F. Wong, "Detect rumors using time series of social context information on microblogging websites," in *Proceedings of the 24th ACM international conference on information and knowledge management*, pp. 1751–1754, Melbourne, Australia, October 2015.
- [25] P. Qi, J. Cao, T. Yang, J. Guo, and J. Li, "Exploiting multi-domain visual information for fake news detection," in *Proceedings of the 2019 IEEE International Conference on Data Mining (ICDM)*, pp. 518–527, IEEE, Beijing, China, November 2019.
- [26] Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo, "Multimodal fusion with recurrent neural networks for rumor detection on microblogs," in *Proceedings of the 25th ACM international conference on Multimedia*, pp. 795–816, Mountain View, CA, USA, October 2017.
- [27] Y. Wang, F. Ma, Z. Jin et al., "Eann: event adversarial neural networks for multi-modal fake news detection," in *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining*, pp. 849–857, London, UK, July 2018.
- [28] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. I. Satoh, "Spotfake: a multi-modal framework for fake news detection," in *Proceedings of the 2019 IEEE fifth international conference on multimedia big data (BigMM)*, pp. 39–47, IEEE, Singapore, September 2019.
- [29] S. Singhal, A. Kabra, M. Sharma, R. R. Shah, T. Chakraborty, and P. Kumaraguru, "Spotfake+: a multimodal framework for fake news detection via transfer learning (student abstract)," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 10, pp. 13915–13916, New York, NY, USA, April 2020.
- [30] L. Cui, S. Wang, and D. Lee, "Same: sentiment-aware multi-modal embedding for detecting fake news," in *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining*, pp. 41–48, Vancouver, Canada, August 2019.
- [31] A. Fukui, D. H. Park, D. Yang, A. Rohrbach, T. Darrell, and M. Rohrbach, "Multimodal compact bilinear pooling for visual question answering and visual grounding," 2016, <https://arxiv.org/abs/1606.01847>.
- [32] W. Yin and H. Schütze, "Multigrannn: an architecture for general matching of text chunks on multiple levels of granularity," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, vol. 1, pp. 63–73, Beijing, China, July 2015.

Research Article

Research on the Construction of College Football Classroom Practice Teaching System Model Based on Big Data Analysis

Bo Zhang¹ and Wei Ren ²

¹PE Department, North Sichuan Medical College, Nan Chong 63700, China

²PE Department of Public Teaching Centre, Cheng Du Medical College, Cheng Du 610500, China

Correspondence should be addressed to Wei Ren; renweity@cmc.edu.cn

Received 7 December 2021; Revised 20 December 2021; Accepted 28 December 2021; Published 14 February 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Bo Zhang and Wei Ren. 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 information technology, the traditional single classroom teaching and passive learning methods of students can no longer meet the needs of all-round development of college students, and its urgent need to integrate with information technology. This article is aimed at the problem of lagging feedback on training results in the traditional teaching model, teachers' active control, students' passive obedience, ignoring the development of students' personality in college football classrooms, and the inability to carry out personalized tracking and quantitative improvement of the training process of students' related abilities. We constructed a college football classroom practice teaching system model based on big data analysis from the perspectives of establishing big data teaching resources, and implementing personalized resource recommendation, optimizing the traditional teaching process, integrating quantitative training, measurement and recording, implementing quantitative intervention, etc. Colleges and universities have carried out experimental observations. Through continuous observation and comparison, it is found that college football classroom practice teaching under big data is more conducive to improving students' football skills and theoretical level than traditional teaching. This model makes full use of the advantages of big data and the combination of practical teaching methods, which can bring students a different learning experience and obtain good teaching effects. It has guiding and reference significance for college football practical teaching.

1. Introduction

In recent years, the field of physical education has made great progress relying on the rapid development of information technology and network technology. However, in the traditional college football classroom teaching evaluation, the results of learning behavior are often obtained through manual recording and measurement, and the efficiency of data record collection, analysis, and graphic and visualization work processing is low; in the teaching form, teachers actively control, and students passively obey, leading to the students' theoretical learning and understanding being not deep enough; in addition, in the practical teaching link, intensive exercises are mainly carried out based on the data results to improve the quality of students' training and lack of attention to the process of training behavior and ignore

the different personalities and characteristics of the students during the training process. Big data is a collection of multiple data carriers. Due to its rapid transmission, and rich and diversified data characteristics, it can perform in-depth mining of collected data. It can better provide real-time data trends for college teachers and students in the classroom practice teaching process. This application will play a very important role in college teachers and students' football teaching practice.

In the context of the era of big data, massive amounts of data are gradually being integrated into modern teaching. How to use these data and information fully and efficiently to better serve sports training and teaching and to further improve the teaching mode and process is currently a hot research topic. Is also one of the problems that educators must solve [1].

At present, teaching workers and scholars have proposed some concepts and methods of using big data to optimize the teaching mode and content of college football. Downer [2], Griffin and Murtagh [3] showed that quantitative teaching can significantly improve students' theoretical reading ability; Gallagher [4], Stromgren et al. [5] conducted quantitative teaching experiments in mathematics teaching, and the results showed that quantitative teaching has a significant effect on students who have difficulty in learning football training. Lindsley [6] introduced what is the era of big data and the typical cases of data application in the era of big data and analyzed why classroom teaching in the era of big data changes from the perspectives of teachers and students and how to respond to the changes; Binder [7] teaches APP courses based on big data, through the experimental process of a large number of student self-study effect analysis data to analyze the 3 types of students in the experiment, according to their process learning data to analyze their learning methods, so as to find students' efficient self-study mode; Griffin [3] aimed at the big data trend, carried out targeted classroom teaching evaluations for college football classrooms, realized the diversification of evaluation subjects, combined formative and summative evaluations, and integrated quantitative and qualitative evaluations; Lambe et al. [8] aimed at the big data trends, put forward the concept of big data classroom, analyzed the characteristics of big data classroom, designed the construction plan of big data classroom, and pointed out the connotation and development trend of big data classroom.

2. Our Works

This article will launch a research on the construction of the college classroom practice teaching system model in the context of big data. Through the proper use of big data, it can bring greater value to the football teaching classroom, so that colleges and universities can obtain more and more detailed information to the greatest extent. The rich data provides decision-making and new ideas for the innovation of teaching practice in colleges and universities in the future.

3. Related Theories

3.1. Behavioral Learning Theory. Skinner [9] is one of the founders of American New Behavioral Psychology. He believes that human behavior is mainly an operational behavior composed of operational reflections, and operational behaviors are behaviors that act on the environment to produce results. Almost all human behaviors are the result of operational enhancement. Football practice teaching can further improve and strengthen football skills through accurate feedback of students' football skills through data. In addition, students can change the reactions of others through the influence of reinforcement. In the context of teaching practice, operational behavior is more representative, so operational reflex is particularly important in the learning process. In 1954, Skinner introduced this theory into teaching, thinking that teaching is to put forward the goals that students should achieve and control the learning

process, supplemented by training, feedback, and corrective remedies, to form the required behavior that is to achieve the goal and give it immediately. Reinforce those behaviors that deviate from the goal or fail to reach the goal; they are corrected without strengthening [10–14].

3.2. Master Learning Theory. The American educational psychologist Bloom put forward the theory of mastering learning for the first time. He believed that mastering learning is to take "all students can learn well" as the guiding ideology, pay attention to the differences between individuals, and adopt various teaching methods to ensure that every student can master the knowledge learned and complete the learning objectives [15–17]. Its basic process is as follows: student orientation-regular teaching-finding mistakes-correcting mistakes-reevaluation.

In the college football classroom practice teaching based on big data analysis, students can learn independently with their own training data and corrective directions and communicate and discuss their doubts with teachers and classmates through the online teaching platform. Teachers can adjust teaching strategies in time through the large amount of data accumulated in the practical teaching process and then gradually complete the teaching goals. This mode is for all students, ensuring that every student can master the knowledge and skills of football theory.

3.3. Training and Measurement. The procedural method of football practice teaching in colleges and universities requires students to train daily and measure their learning effects, that is, to spend a certain amount of time every day for football training and measurement. Football training is the basis of measurement, and this process requires long-term development and continuous recording. Generally speaking, the measured frequency data will be recorded by the teacher in a standard variable speed chart. The chart can accurately predict when the student's knowledge or skills will meet the training requirements to determine whether the current student's learning performance has improved significantly over time. Whether the intervention measures need to be modified. In the teaching process, the teacher can collect corresponding data according to the students' daily football training, measurement and recording, and frequently monitoring the student's training behavior status to obtain the student's development [18–21].

4. Construction of College Football Practice Teaching System Model Based on Big Data Analysis

In the traditional teaching environment, teachers are more inclined to apply a certain mature teaching model, and practical teaching under big data analysis is often regarded as a teaching evaluation strategy or a teaching method for a certain course rather than a teaching model. Practical teaching under big data analysis first encountered the obstacles of teachers' thinking and ideas in the application and promotion. Big data breaks through many constraints of the

traditional teaching environment and is conducive to encouraging teachers to accept and recognize practical teaching under big data analysis in terms of thinking. It is of great significance to promote the development of practical teaching under big data analysis and to promote the application of practical teaching under big data analysis. For this reason, this research constructs a practical teaching model based on big data analysis based on big data from the three dimensions of teaching goal establishment, teaching process frame design, teaching evaluation, and prediction, as shown in Figure 1.

4.1. Establishment of Practical Teaching Goals in Football Classrooms in Colleges and Universities. The clear goal of practical teaching is the logical starting point for the implementation of teaching, and it is also an important basis for testing the success or failure of teaching. Accordingly, the primary task of practical teaching under big data analysis is to establish practical teaching goals. In the traditional teaching environment, the teaching goal can be ambiguous. For example, in the basic computer course, the teaching goal of a certain course can be “proficient in the conversion of decimal and binary,” and the “mastery” is a fuzzy Degree word. However, in practical teaching under big data analysis, quantifiable teaching goals must be designed; that is, there must be a quantitative explanation and description of the degree of knowledge or skills mastered by students; the basic idea of explanation is the decomposition and refinement of the problem. The way of description is quantification. In other words, in practical teaching under big data analysis, each teaching objective should be transformed into a corresponding problem, and each problem should be decomposed and refined into small problems that can be quantified and described. For example, “proficient in the conversion of decimal and binary” can be transformed into “completing 5 decimal and binary exchange problems within 1000 within 3 minutes, and the accuracy rate is 100%”; the “proficient” here has been decomposed and refined. After quantification, it includes not only the accurate mastery of knowledge or skills, but also the speed of applying knowledge or skills, so as to keep fully in line with the measurement indicators of practical teaching under big data analysis.

4.2. The Framework Design of Programmatic Teaching Process in College Football Class. Practical teaching under big data analysis originated from Skinner’s program teaching, so proceduralization is the core element of practical teaching under big data analysis. Designing a procedural teaching process framework is the key to ensuring the effective implementation of practical teaching under big data analysis. The programmatic teaching process framework designed in this research refers to the process and rules of practical teaching based on big data implementation under big data analysis, including:

4.2.1. Establish a Big Data Teaching Resource Database and Implement Personalized Resource Recommendation. The essence of programmed teaching is an input and output

system; that is, input teaching resources and output the results of football training of students. In the traditional teaching environment, due to limited teaching resources and lack of information technology, the input and output system takes the entire teaching class as the basic particle, which cannot guarantee the individualized development of students. In response to this problem, this research proposes to establish a big data education resource library to manage massive digital teaching resources; at the same time, the basic particles of the input and output system are refined from the class to each student, using intelligent recommendation technology, according to the students. The characteristics of learning are equipped with different high-quality teaching resources, and personalized teaching is implemented.

4.2.2. Optimize the Traditional Teaching Process and Incorporate Quantitative Training, Measurement and Recording. This research makes full use of the advantages of big data technology to optimize the traditional teaching process, on this basis, incorporates quantitative training, measurement, and recording, and conducts practical teaching under big data analysis, so as to provide support for the next step of teaching decision-making and learning intervention. Specifically, this research has conducted two types of attempts at the operational level:

- ① Based on excellent online resources to assist teaching, that is, using micro-classes as the content and youtube as the platform, practical teaching under big data analysis can fully meet the teaching needs-first, and students pay attention to the youtube course number with their real names; then, students click on the youtube platform Online micro-class resources, theoretical study, and offline training; finally, the background system automatically records the student’s learning behavior to form each student’s theoretical learning trajectory and analysis results.
- ② Practical teaching based on project-oriented task-driven, that is, based on the computer basic course training and evaluation system as a platform, in the traditional project-oriented task-driven teaching framework, practical teaching under big data analysis is carried out-first, students log in to the system for practical training Drill, and you can submit every task you complete; otherwise, you cannot enter the next task stage; after the entire project is completed, submit it to the system for scoring; finally, the system records the time and points lost for students logging in and completing each training task in real-time (Error) and the final score form the student’s learning trajectory and error problem domain.

4.2.3. Implement Quantitative Interventions. Quantitative intervention is the essence of practical teaching under big data analysis. In a big data environment, whether it is

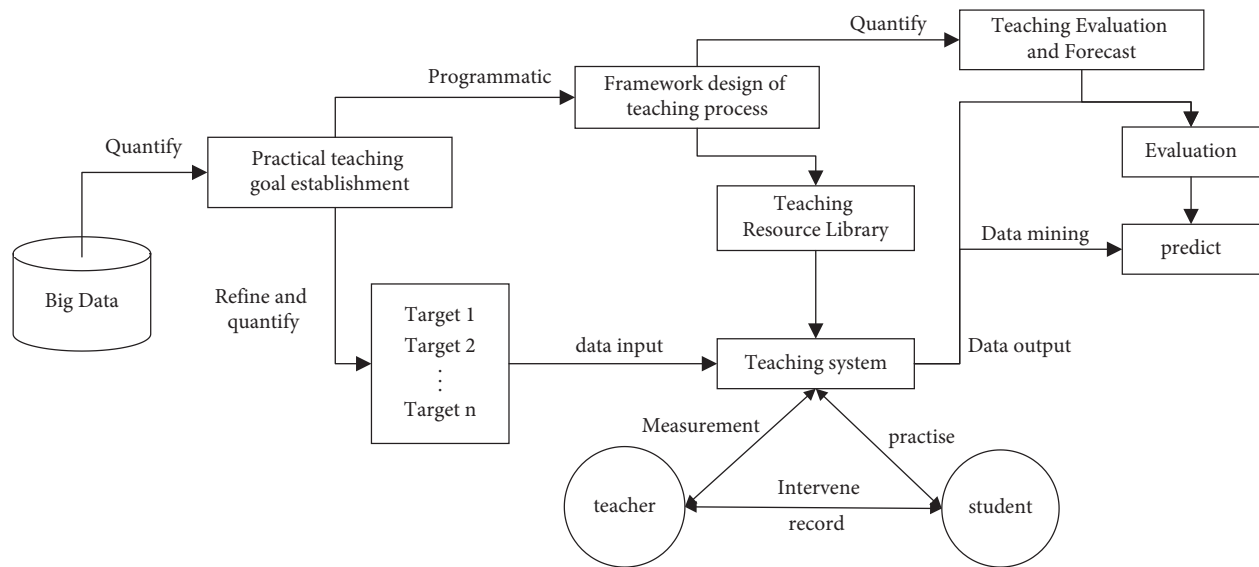


FIGURE 1: University football classroom practice teaching system model based on big data analysis.

youtube or a computer basic course practice evaluation system, teachers and students can communicate across time and space, and communication records can be traced back and forth. Based on the measured and recorded football training behaviors of students, teachers can judge whether students can successfully achieve the teaching goals—if they can be achieved, there is no problem; if they cannot be achieved, there is a problem, and intervention is required. Specifically, this research carried out targeted interventions at the operational level according to special problems and general problems: for the special problems of individual students, real-time point-to-point intervention and correction are carried out through instant messaging tools; for common problems that are more reflected, they are unified intervention and correction through teaching blogs, youtube, and Moments. Intervention is an iterative task, and football training, measurement, and recording are combined with intervention. In each stage of the training process, the effectiveness of the training is tracked and feedback point-to-point, and it becomes a cyclic iterative process. This cyclic iteration will not stop until all students have reached the theoretical knowledge of football or football skills required by the teaching goal.

4.3. Evaluation and Forecast of Practical Teaching in Football Class in Colleges and Universities. In the traditional teaching environment, teaching evaluation may be a vague empirical judgment, such as using degree words such as “excellent,” “good,” “medium,” “pass,” and “poor” to evaluate students’ learning performance; or simple Score judgment, such as the final exam score, mid-term exam score, total score, and average score to evaluate the student’s learning results. In the big data environment, the fusion application of many advanced technologies such as sensor technology, face recognition technology, and learning analysis technology makes the evaluation of practical teaching under big data analysis from the beginning to the end of the accompanying

teaching behavior and can be used to assess what has not happened yet. Quantitative forecasts for the future. For example, Zheng Yiwen et al. proposed a classroom big data collection technology, which integrates student sitting posture measurement system, eye recognition system, and noise recognition system. By acquiring some big data on students’ living conditions in the classroom, it can be more accurate Interpretation, analysis, and judgment of students’ learning conditions (such as class arrival, concentration of thoughts, classroom activity, and physical fatigue); this technology has high real-time performance, making it possible to implement quantitative and effective attention to each student possible. It can be seen that the practical teaching evaluation based on big data analysis is a kind of real-time evaluation of the whole staff, the whole process, and all directions. In the practical teaching model based on big data analysis, teaching evaluation mainly relies on technical means (including big data collection, educational data mining, learning analysis, and data visualization technology) and is automatically monitored and automatically monitored by various intelligent teaching systems. Analyze the learning situation of students and give real-time feedback to those who need it; teachers, students, parents, etc. can query and generate visual evaluation reports according to their own needs. Prediction refers to the comprehensive analysis of each student’s learning performance data and other system data (including various education systems, evaluation systems, and expert systems) to form a data decision support system and conduct a comprehensive analysis of the student’s learning performance in the future. Forecast, and then put forward relevant football training improvement suggestions or countermeasures based on the forecast results.

4.4. Selection and Evaluation of Experimental Indexes in College Football Class. A large number of practical studies

have shown that the use of big data analysis in physical education can better improve students' physical fitness and technical level, cultivate students' sensitivity to training quantitative results, and exercise students' ability to independently develop quantitative results. Compared with the traditional teaching mode, its effect is more significant. Therefore, in order to further study it, to explore whether the application of this model in college football classroom practice teaching can have a positive impact. Based on big data analysis of college football practice teaching methods and combined with previous research results, the following core weight data indicators recorded in the teaching system in real time in the past are selected to carry out experimental research in a college:

4.4.1. Physical Fitness. Physical fitness refers to the strength, speed, endurance, agility, and other functional capabilities of the human body in activities and is the basis for mastering sports techniques and improving sports performance. After quantifying and screening through big data, scoring is performed, and the top three items with the highest weight coefficients are selected as the test indicators of physical fitness in this experiment. After screening, the 5 * 25-meter retracement run, the 30-meter sprint, and the standing long jump are used as the physical fitness test indicators (see Table 1 for details).

Before and after the experiment, the two groups of students were tested for their physical fitness. The test content included three indicators: 30-meter sprint, standing long jump, and 5 * 25-meter return run. The 30-meter sprint can reflect the quality of speed, the standing long jump can reflect the explosive power of the lower limbs, and the 5 * 25-meter return run can reflect the physical quality such as speed, endurance, and agility. Figure 2 shows the test site, where the experimental group performed a 5 * 25-meter turnback run.

4.4.2. Football Technology. According to the technical assessment requirements of efficient football practice teaching, and consulting football related teachers, determine the frontal bump and shot of the instep as the football technical test indicators. The assessment content and scoring standards are as follows:

① Bump the ball on the front of the instep

Test method: the teacher starts to issue the order after the tester is ready. The tester picks up the ball with his toes or throws the ball with his hands and then uses the front of the instep to make continuous ball bumps. During the ball bumping process, the tester tries his best to turn his feet alternately. The ball can also be adjusted with one foot continuously, and the test is deemed to be terminated once the ball hits the ground. Each person has three test opportunities to get the best result. The scoring standards for the ball-dumping are shown in Table 2:

② Shot

Test method: boys shoot from the penalty area line, girls shoot from the penalty spot, divide the goal into three equals, shoot two points on both sides, and score one point in the middle. Each person has five chances and accumulates points. Figure 3 shows the test site of the football shooting technique of the experimental group.

4.4.3. Theoretical Assessment. It is proposed to select certain college football theory assessment regulations. There are three sets of theoretical examination papers ABC. One week before the examination, teachers randomly select one of them as the final examination papers. Take the three knowledge modules of football overview, basic football skills and tactics, and football game rules as the theoretical investigation content. It mainly includes four question types: judgment, filling in the blanks, noun explanation, and question and answer. There are 22 questions in total, with a total score of 100.

5. Experimental Results and Analysis

5.1. The Results and Analysis of Various Indicators before the Experiment in the Experimental Group and the Control Group

5.1.1. Detection and Recognition Effect. Good physical fitness is the foundation and guarantee of football, and it is also an important part of football practice teaching effect. This chapter will use the experimental group and the control group to carry out the experimental research of college football classroom practice teaching mode under big data. First, in order to test whether there is a difference in the physical fitness of the students, the students in the two classes were tested for 30 meters, 5 * 25 reentry running, and standing long jump before the experiment began, and the two classes were premeasured with big data. Analyze the independent sample *T* test, and the results are as follows:

From Table 3, it can be concluded that the 30-meter score of the experimental group is 4.40 ± 0.08 s, the 5 * 25 foldback running score is 33.79 ± 1.10 s, and the standing long jump is 2.67 ± 0.08 m; the 30-meter score of the control group is 4.37 ± 0.06 , 5 * 25 foldback running result was 33.09 ± 1.27 s, and the standing long jump was 2.70 ± 0.07 m. In order to see more clearly whether there are differences in the physical fitness of the two classes, an independent sample *T* test was performed on the results of the two classes, and the results showed that $P_{30m} = 0.083 > 0.05$, $P_{Turn\ back} = 0.072 > 0.05$, $P_{standing\ long\ jump} = 0.13 > 0.05$. It can be seen that there is no difference in physical fitness between the experimental group and the control group, which meets the requirements of the experiment.

5.1.2. Pretest Results and Analysis of Football Skills. In accordance with the test subject's college football assessment standards, the experimental group and the control group were tested on two indicators of ball bumping and shooting before the experiment. The two indicators are subjected to

TABLE 1: Quantitative screening scores of physical fitness test indicators.

Physical fitness test index	Influence level					Total score
	Very important 5 points	Important 4 points	Generally 3 points	Not so important 2 points	Unimportant 1 points	
30-meter sprint	4	2	0	0	0	28
12 minutes run	1	1	0	2	2	15
Standing long jump	3	2	1	0	0	26
YOYO	1	2	0	3	0	19
1000 m	0	0	3	2	1	14
Back and forth	5	1	0	0	0	29

Note. Perform linear interpolation scoring based on the quantitative results.



FIGURE 2: Turn-back running test.

TABLE 2: Score and score calculation table.

Score	Male	3	4	5	6	8	10	15	20	25
	Female	2	3	4	5	6	8	10	15	20
Points		2	3	4	5	6	7	8	9	10



FIGURE 3: Football technical assessment.

independent sample T test under big data analysis, and the results are as follows:

From Table 4 above, it can be concluded that the experimental group's smashing score was 2.35 ± 0.93 points, and the shooting score was 4.05 ± 0.95 points; the control group's smashing score was 2.75 ± 1.12 points, and the shooting score was 4.25 ± 1.21 points. Clearly see whether there is a difference in the physical fitness of the two classes.

The results of the two classes are subjected to an independent sample T test under big data analysis, and the result is that $P_{\text{Bumping the ball}} = 0.227 > 0.05$, $P_{\text{shot}} = 0.563 > 0.05$, and it can be seen that there is no significant difference between the experimental group and the control group in the two techniques of throwing the ball and shooting. The football skills of the two classes are basically at the same level, which guarantees football teaching accuracy of teaching experiments.

Through the pretest, I have a basic understanding of the football skills of the students in the two classes. The test results reflect that the two classes have poor ball and shooting skills. Only a small number of students have good skills. According to the assessment standards, most students have a good performance. Being at a failing level, this shows that the students have not received professional and systematic football practice teaching before, and rarely do football sports after class.

5.2. *The Results and Analysis of Each Index after the Experiment of the Experimental Group and the Control Group.* After 16 weeks of continuous follow-up of practical teaching, the two classes of students' physical fitness, football technical level, and football theory learning three aspects were once again compared and analyzed. According to the syllabus of the subject colleges and universities, football skills accounted for 50%, football theory 30%, and football basic technology teaching 20% in the final assessment standards of football. For this reason, this article adds football theory and football

TABLE 3: Premeasurement results of physical fitness of the experimental group and the control group.

Test index	Test group ($\bar{X} \pm S$) ($N=20$)	Control group ($\bar{X} \pm S$) ($N=20$)	T value	P value
30 m (s)	4.40 \pm 0.08	4.37 \pm 0.06	1.782	0.083
Turn back (s)	33.79 \pm 1.10	33.09 \pm 1.27	1.848	0.072
Standing long jump (m)	2.67 \pm 0.08	2.70 \pm 0.07	-1.547	0.130

TABLE 4: Premeasurement results of football skills of the experimental group and the control group.

Test content	Test group ($\bar{X} \pm S$) ($N=20$)	Control group ($\bar{X} \pm S$) ($N=20$)	T value	P value
Bumping the ball (points)	2.35 \pm 0.93	2.75 \pm 1.12	-1.228	0.227
Shots (points)	4.05 \pm 0.95	4.25 \pm 1.21	-0.583	0.563

basic technology teaching in the postexperimental test. Item test indicators were used in order to more comprehensively judge the impact of the college football classroom practice teaching system mode on the football teaching effect.

5.2.1. Pretest Results and Analysis of Physical Fitness

(1) *Comparison of Physical Fitness between the Experimental Group and the Control Group before and after the Experiment.* In the last week of the experiment, the three indicators of the long jump were tested. The same as the previous test, the 30-meter run, 5 * 25 reentry run, and standing of the two classes of students were tested to test whether the football practice teaching system model under big data has an impact on the physical fitness of students. The comparative data of the physical fitness of the experimental group and the control group before and after the experiment are as follows:

Table 5 shows that, after the experiment, the three indicators of the physical fitness of the experimental group before and after the experiment were, respectively, subjected to the paired sample T test under the big data analysis, and the result was $P_{30m} = 0.083 > 0.05$, $P_{Turn\ back} = 0.475 > 0.05$, $P_{standing\ long\ jump} = 0.087 > 0.05$. It can be seen that, after the experiment, the physical fitness of the students in the experimental group does not change significantly, and there is no significant difference.

Table 6 shows that, after the experiment, the three indicators of the physical fitness of the control group before and after the experiment were, respectively, subjected to the paired sample T test under big data analysis, and the result was $P_{30m} = 0.175 > 0.05$, $P_{Turn\ back} = 0.071 > 0.05$, $P_{standing\ long\ jump} = 0.385 > 0.05$. It can be seen from this that, before and after the experiment, the three indicators of physical fitness of the control group students did not change significantly.

(2) *Comparison of Physical Fitness between the Experimental Group and the Control Group after the Experiment.* Table 7 shows that, after the experiment, the three physical fitness scores of the two classes of students were subjected to an independent sample T test under big data analysis. The results showed that the P values of the three physical fitness items were all greater than 0.05. The experimental group and the control group were posttested. The physical fitness is not much different, and there is no significant difference.

(3) *Analysis of Physical Fitness Test Results.* The above results show that, after the experiment, the physical fitness of the students in the experimental group has not significantly improved, and there is no significant difference between the physical fitness scores of the experimental group and the control group. The objective reasons above are mainly due to the following two points:

- (1) The human body is a gradual adaptation to sports stimulation, a cyclical process. Students' physical fitness and physical functions need to develop gradually and slowly from quantitative changes to qualitative changes in a continuous environment [22, 23]. In the 16-week college football classroom practice teaching under big data, if you want to improve the students' physical fitness, you must not only arrange the load intensity and load reasonably, but also follow the "law of human body function adaptability," even if next time the load is arranged in the overload recovery stage of the last load, forming a relatively stable state of load-adaptation-additional load-readaptation (Figure 4) [24]. According to the schedule of football training sessions, the interval between the two sessions is too long. The load and recovery between classes cannot form an effective connection. Therefore, the physical fitness of students is difficult to significantly improve under various conditions.
- (2) From Tables 5 and 6, it can be seen that after, 16 weeks of football practical teaching under big data, the preexperimental results of the 5 * 25 m reentry run of the experimental group and the control group were significantly better than the postexperimental results. The reason for the change of this indicator may be because the measured time after the experiment was December. The climate is cold in winter, and the human body will reflexively cause muscle and blood vessel contraction under cold stimulation, which will reduce the elasticity of muscles and ligaments, the joints are small, and the students' physical activity is not sufficient. The degree of excitement is not high, so the measured results can hardly reflect the true level of the students. The law of human body function adaptability has no essential causal relationship with data

TABLE 5: Comparison and quantitative results of the physical fitness of the experimental group before and after the experiment.

Test index	Before the experiment ($\bar{X} \pm S$) ($N=20$)	After the experiment ($\bar{X} \pm S$) ($N=20$)	T value	P value
30 m (s)	4.40 \pm 0.08	4.38 \pm 0.05	1.833	0.083
Turn back (s)	33.79 \pm 1.10	34.02 \pm 0.93	-0.728	0.475
Standing long jump (m)	2.67 \pm 0.08	2.68 \pm 0.07	-1.807	0.087

TABLE 6: Comparative quantitative results of physical fitness before and after the experiment of the control group.

Test index	Before the experiment ($\bar{X} \pm S$) ($N=20$)	After the experiment ($\bar{X} \pm S$) ($N=20$)	T value	P value
30 m (s)	4.37 \pm 0.06	4.38 \pm 0.05	-1.407	0.175
Turn back (s)	33.09 \pm 1.27	33.58 \pm 0.74	-1.916	0.071
Standing long jump (m)	2.70 \pm 0.07	2.71 \pm 0.07	-0.890	0.385

TABLE 7: Quantified comparison results of physical fitness between the experimental group and the control group after the experiment.

Test index	Before the experiment ($\bar{X} \pm S$) ($N=20$)	After the experiment ($\bar{X} \pm S$) ($N=20$)	T value	P value
30 m (s)	4.38 \pm 0.05	4.38 \pm 0.05	-0.271	0.788
Turn back (s)	34.02 \pm 0.93	33.58 \pm 0.74	1.656	0.106
Standing long jump (m)	2.68 \pm 0.07	2.71 \pm 0.07	-1.394	0.107

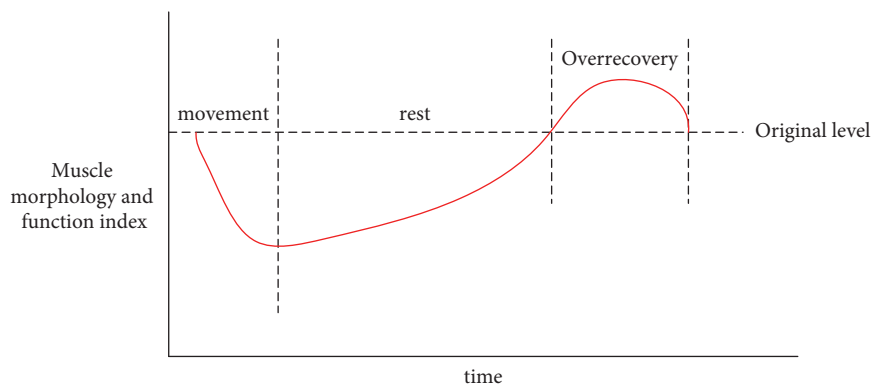


FIGURE 4: The law of human body function adaptability.

analysis, but it will have an impact on the results. In addition, the physical fitness of students is affected by physical conditions, environmental weather, and other factors. It is difficult to improve the physical fitness of students through this experiment. The focus of the football practice teaching system model under big data lies in cultivating students' autonomous learning ability and training result-oriented ability rather than improving students' physical fitness.

5.2.2. Results and Analysis of the Posttest of Football Skills

(1) *The Comparison Result of Football Skill Level Test between the Experimental Group and the Control Group before and after the Experiment.* The level of football skills is an important indicator that reflects the implementation effect of the college football classroom practice teaching system model and the traditional classroom model under the big

data. After 16 weeks of experimental observation, in order to test whether the college football classroom practice teaching system model under big data has a positive impact on students' football skills in football lessons, the experimental group and the control group football technical pre-and posttest scores are compared and analyzed with big data. The comparison results are as follows:

Tables 8 and 9 show that, through the paired-sample T test under the large number analysis of the scores of the experimental group and the control group before and after the experiment, the test results of the two classes after the experiment are higher than those of the experiment. The pretest results show that both the college football classroom practice teaching system model and the traditional teaching model under the big data can improve students' football skills.

(2) *The Comparison Result of Football Skill Level Test between the Experimental Group and the Control Group after the Experiment.* It can be seen from Table 10 that a comparative

TABLE 8: Results of quantitative comparison of football skills before and after the experiment in the experimental group.

Test index	Before the experiment ($\bar{X} \pm S$) ($N=20$)	After the experiment ($\bar{X} \pm S$) ($N=20$)	T value	P value
Bumping the ball (points)	2.35 ± 0.93	6.95 ± 1.73	-13.707	$P < 0.001$
Shots (points)	4.05 ± 0.95	7.50 ± 1.24	-11.716	$P < 0.001$

TABLE 9: Quantitative comparison results of football skills before and after the experiment of the control group.

Test index	Before the experiment ($\bar{X} \pm S$) ($N=20$)	After the experiment ($\bar{X} \pm S$) ($N=20$)	T value	P value
Bumping the ball (points)	2.75 ± 1.12	5.25 ± 1.02	-11.180	$P < 0.001$
Shots (points)	4.25 ± 1.21	6.25 ± 1.12	-6.325	$P < 0.001$

TABLE 10: Quantified comparison results of football skills between the experimental group and the control group after the experiment.

Test index	Before the experiment ($\bar{X} \pm S$) ($N=20$)	After the experiment ($\bar{X} \pm S$) ($N=20$)	T value	P value
Bumping the ball (points)	6.95 ± 1.73	5.25 ± 1.02	3.784	$P < 0.001$
Shots (points)	7.50 ± 1.24	6.25 ± 1.12	3.355	$P < 0.001$

analysis of the scores of the ball and shooting of the experimental group and the control group shows that the two posttest scores of the experimental group are higher than those of the control group, and there is a very significant difference ($P < 0.01$).

(3) *Analysis of the Test Results of Football Skills.* According to the above results, after 16 weeks of practical teaching experiment observation, the football technical performance of the experimental group and the control group have improved, but the football technical level of the experimental group has improved more significantly. The analysis of the reasons is mainly in the following two aspects:

- (1) The comparative analysis of the pre- and posttest scores of the experimental group and the control group can be obtained: after 16 weeks of practical teaching experiment, under the big data college football classroom practical teaching system model and the traditional teaching model, the two classes of students level of ball and shooting skills have been significantly improved, which shows that the college football classroom practice teaching system model under big data still has a certain effect on the improvement of students' football skills, and the technical skill training for students is still very advantageous.
- (2) The comparative analysis of the football technical posttest scores of the experimental group and the control group can be obtained: after the experiment, the posttest scores of ball throws and shots of the experimental group are higher than those of the control group, and there are very significant differences. This shows that, in improving students' football skills, the implementation effect of the college football classroom practice teaching system model under big data is significantly better than the traditional classroom teaching model. Analyzed from the perspective of the law of

formation of motor skills, students learn the teaching content pushed by teachers independently and correct training actions in time based on big data information feedback to form motor skills. The time for teachers to explain and demonstrate in class will be greatly shortened, and students will have enough time to reinforce exercises and practice. Teachers organize students' exercises and carry out differentiated error correction and personalized guidance through teachers' feedback and peer interaction, learning, discussing, abandoning the cognition of wrong movements, perfecting and optimizing one's own technical movements, so that the development of football skills can be improved through continuous feedback and correction. Due to the timely feedback of training data, students become more motivated to learn and have a stronger desire to achieve goals. Therefore, using the college football classroom practice teaching system model under big data for football teaching can improve the students' football skills, and its teaching effect is better than the traditional teaching model.

5.2.3. The Results and Analysis of the Posttest of Football Theoretical Results

(1) *The Comparison Result of the Experimental Group and the Control Group after the Experiment of Football Theory Score Test.* Football projects focus on the combination of theory and practice, and mastering theoretical knowledge can greatly promote the learning of football skills. Therefore, the degree of students' mastery of theory is a key indicator reflecting the effectiveness of college football classroom practice teaching under big data. After the 16-week experimental observation, a posttest comparative analysis of the theoretical football scores of the experimental group and the control group was carried out. The results are as follows:

TABLE 11: Comparison results of theoretical football scores between the experimental group and the control group after the experiment.

Content	Test group ($\bar{X} \pm S$) ($N=20$)	Control group ($\bar{X} \pm S$) ($N=20$)	T value	P value
Football theory score (100 points)	88.10 \pm 4.58	76.15 \pm 8.60	5.484	$P < 0.001$

As can be seen from the above table, after the experiment, the theoretical scores of the two classes of students are processed by independent sample T test under big data. The results show that the theoretical scores of the experimental group are 88.10 ± 4.58 points, and the theoretical scores of the control group are 76.15 ± 8.60 points. $P < 0.01$, the theoretical posttest scores of the experimental group are significantly higher than those of the control group, and there are very obvious differences.

(2) *Analysis of the Test Results of Football Theory.* It can be seen from Table 11 that, after the experiment, there is a significant difference between the experimental group and the control group in football theoretical performance, indicating that college football classroom practice teaching under big data is more conducive to improving students' grasp of football theoretical knowledge than traditional teaching. A detailed analysis of the reason is mainly due to the fact that college football classroom practice teaching under big data is goal- and result-oriented, programmatic learning and effective feedback can be carried out, and the feedback of big data information can be used for timely evaluation and further strengthening of learning during the learning process. In addition, in the classroom, the teacher can use the results of the systematic evaluation and recording to further intervene in the learning situation of the students. In order to improve the level of football skills in the learning and practice, the students will actively think about the technical movements and strengthen the theory of football while mastering football skills, understanding and deepening of knowledge.

6. Conclusion

This paper studies the college football classroom practice teaching system model based on big data analysis and carries out a procedural design for the college football classroom teaching process. Under the big data analysis, the corresponding evaluation indicators are selected according to the weight of the indicators to carry out experimental comparative research. According to the continuous collection of data in the teaching practice, the following conclusions are drawn: (1) the experimental group and the control group under the big data analysis have little difference in physical fitness after testing, and there is no significant difference. (2) Fully adopting the college football classroom practice teaching system model under big data to carry out football teaching can improve the football skills of students, and its teaching effect is better than traditional teaching mode. (3) College football classroom practice teaching under big data focuses on goal and result orientation and can carry out procedural learning and effective feedback. In the learning

process, the feedback of big data system information can be evaluated in time and further strengthened to strengthen the theory.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was supported by the Special Project of Science and Technology Strategic Cooperation of municipal schools: Research on management and evaluation system of physical health promotion for college students (20SXGJSK0004); and Research Center of Yi Culture Si Chuan (YZWH1603).

References

- [1] X. Yang, L. Wang, and D. Sisi, "The application model and policy suggestion of big data in education," *Audio-visual Education Research*, vol. 23, no. 9, pp. 54–61, 2015.
- [2] A. C. Downer, "The national literacy strategy sight recognition programme implemented by teaching assistants: a precision teaching approach," *Educational Psychology in Practice*, vol. 23, no. 2, pp. 129–143, 2007.
- [3] C. P. Griffin and L. Murtagh, "Increasing the sight vocabulary and reading fluency of children requiring reading support: the use of a precision teaching approach," *Educational Psychology in Practice*, vol. 31, no. 2, pp. 186–209, 2015.
- [4] E. Gallagher, "Improving a mathematical key skill using precision teaching," *Irish Educational Studies*, vol. 25, no. 3, pp. 303–319, 2006.
- [5] B. Strømgren, C. Berg-Mortensen, and L. Tangen, "The use of precision teaching to teach basic math facts," *European Journal of Behavior Analysis*, vol. 15, no. 2, pp. 225–240, 2014.
- [6] O. R. Lindsley, "Precision teaching: discoveries and effects," *Journal of Applied Behavior Analysis*, vol. 25, no. 1, pp. 51–57, 1992.
- [7] C. Binder, "Precision teaching: measuring and attaining exemplary academic achievement," *Youth Policy*, vol. 31, no. 7, pp. 12–15, 1988.
- [8] D. Lambe, C. Murphy, and M. E. Kelly, "The impact of a precision teaching intervention on the reading fluency of typically developing children," *Behavioral Interventions*, vol. 30, no. 4, pp. 364–377, 2015.
- [9] B. F. Skinner, *Science and Human Behavior*, pp. 45–140, Macmillan, New York, NY, USA, 1953.
- [10] P. Stansbie, R. Nash, and S. Chang, "Linking internships and classroom learning: a case study examination of hospitality and tourism management students," *Journal of Hospitality, Leisure, Sports and Tourism Education*, vol. 19, pp. 19–29, 2016.

- [11] D. Sledgianowski, M. Gomaa, and C. Tan, "Toward integration of Big Data, technology and information systems competencies into the accounting curriculum," *Journal of Accounting Education*, vol. 38, pp. 81–93, 2017.
- [12] M. Shorfuzzaman, M. S. Hossain, A. Nazir, G. Muhammad, and A. Alamri, "Harnessing the power of big data analytics in the cloud to support learning analytics in mobile learning environment," *Computers in Human Behavior*, vol. 92, pp. 578–588, 2019.
- [13] M. Cantabella, R. Martínez-España, B. Ayuso, J. A. Yáñez, and A. Muñoz, "Analysis of student behavior in learning management systems through a Big Data framework," *Future Generation Computer Systems*, vol. 90, pp. 262–272, 2019.
- [14] R. R. Hoy, "Quantitative skills in undergraduate neuroscience education in the age of big data," *Neuroscience Letters*, vol. 759, Article ID 136074, 2021.
- [15] Y. Jiang, L. Ma, and L. Gao, "Assessing teachers' metacognition in teaching: the teacher metacognition inventory," *Teaching and Teacher Education*, vol. 59, pp. 403–413, 2016.
- [16] X. Bai, F. Zhang, J. Li et al., "Educational big data: predictions, applications and challenges," *Big Data Research*, vol. 26, Article ID 100270, 2021.
- [17] D. Sampaio and P. Almeida, "Pedagogical strategies for the integration of augmented reality in ICT teaching and learning processes," *Procedia Computer Science*, vol. 100, pp. 894–899, 2016.
- [18] X. He, K. Yu, Z. Huang et al., "Multilevel-teaching/training practice on GNSS principle and application for undergraduate educations: a case study in China," *Advances in Space Research*, vol. 69, 2021.
- [19] P. Xiao-Pang, G. N. Vivekananda, and S. Khapre, "Multi-media-based English teaching and practical system," *Aggression and Violent Behavior*, vol. 25, no. 1, Article ID 101706, pp. 51–57, 2021.
- [20] B. Dietz-Uhler and J. R. Lanter, "Using the four-questions technique to enhance learning," *Teaching of Psychology*, vol. 36, no. 1, pp. 21–24, 2009.
- [21] F. J. Hinojo-Lucena, A. C. Mingorance-Estrada, J. M. Trujillo-Tomes, I. Aznar-Diaz, and M. P. C. Reche, "Incidence of the flipped classroom in the physical education Students' Academic performance in university contexts," *Sustainability*, vol. 10, no. 5, 2018.
- [22] G. C. Gannod, J. E. Surge, and M. T. Helmick, "Using the inverted classroom to teach software engineering," in *Proceedings of the ACM/IEEE 30th International Conference on Software Engineering*, pp. 777–786, Leipzig, Germany, May 2008.
- [23] J. W. Baker, "The classroom flip using web course management tools to become the guide by the side," in *Proceedings of the Selected papers from the 11th International Conference on College Teaching and Learning*, Jacksonville, Florida, April 2000.
- [24] W. Welling, A. Benjaminse, K. Lemmink, B. Dingenen, and A. Gokeler, "Progressive strength training restores quadriceps and hamstring muscle strength within 7 months after ACL reconstruction in amateur male soccer players," *Physical Therapy in Sport*, vol. 40, pp. 10–18, 2019.

Retraction

Retracted: Visual Evaluation of Urban Streetscape Design Supported by Multisource Data and Deep Learning

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

References

- [1] G. Feng, G. Zou, and P. Wang, "Visual Evaluation of Urban Streetscape Design Supported by Multisource Data and Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3287117, 9 pages, 2022.

Research Article

Visual Evaluation of Urban Streetscape Design Supported by Multisource Data and Deep Learning

Guanqing Feng,^{1,2,3} Guangtian Zou ^{1,2} and Pengjin Wang³

¹Department of Architecture, Faculty of Architecture, Harbin Institute of Technology, Harbin 150006, Heilongjiang, China

²Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology (Harbin Institute of Technology), Harbin 150006, Heilongjiang, China

³Department of Landscape Architecture, Faculty of Horticulture and Landscape Architecture, Northeast Agricultural University, Harbin 150030, Heilongjiang, China

Correspondence should be addressed to Guangtian Zou; fengguanqing@neau.edu.cn

Received 23 November 2021; Accepted 13 January 2022; Published 7 February 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Guanqing Feng 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 integrates classical design theory, multisource urban data, and deep learning to explore an accurate analytical framework in a new data environment, providing a scientific analysis path for the “where” and “how” of greenways in a high-density built environment. The analysis is based on street view data and location service data. Through the integration of multiple data sources such as street scape data, location service data, point-of-interest data, structured web data, and refined built environment data, a systematic measurement of the key elements of density, diversity, design, accessibility to destinations, and distance to transport facilities as defined in the Five Elements of High Quality Built Environment (5D) theory is achieved. The assessment of alignment potential was carried out. The key factors influencing the aesthetics of the street were identified. Based on an extensive landscape perception-based survey, it was found that although different respondents had different views and preferences for the same street scape, their preferences were overwhelmingly influenced by the visual quality of the street scape aesthetics itself, with higher aesthetic quality of the landscape.

1. Introduction

In the landscape engineering system, landscape resources are divided into 3 categories: human resources, natural resources, and landscape visual resources [1]. Of these, how to evaluate landscape resources, especially landscape aesthetics, has not yet been agreed upon by the landscape community and is expressed in practice in inconsistent ways [2]. Most people will own their own opinions on whether a particular landscape is aesthetically significant, and there is an increasing emphasis on the role that everyday landscapes play in everyday life [3]. In the case of landscape aesthetics, there are many controversies in the industry. Often, the ambiguity of landscape aesthetics is due to the subjectivity of its definition, the lack of standardisation of its methodology, the lack of clarity in its practical application, and its own

particular lack of replicability [4]. In contrast, it has become increasingly evident in the last decades that the visual aesthetic quality of landscapes is considered an important resource for the maintenance of people’s psychological well-being, as well as for the conservation of biodiversity, cultural heritage, and the potential of landscapes [5–8]. Therefore, the development of a methodology for the analysis and evaluation of landscape aesthetics that can be accepted by the scientific community as a whole is a major challenge facing the academic community today.

The visual aesthetic quality can be viewed from the same 2 perspectives. From an objective point of view, the visual aesthetic quality of a landscape is determined by the elements it contains and its properties. The subjective point of view, on the other hand, explores the status of the visual aesthetics of the landscape within people and the subjective

feelings it brings to them, starting from their subjective inner thoughts. Paper [9] has also proposed in the systematisation of landscape engineering that the intentional content of the landscape carries objective factors and the real content of the landscape carries subjective factors and that the aesthetic feeling of the landscape always contains these 2 types of basic components, while the controlling role is played by the intentional content of the landscape. Based on the above background, two approaches to assessing the visual aesthetics of landscapes are commonly used in the industry [10].

In addition, certain specific visual attributes of the landscape are also considered important, such as openness, consistency, colour contrast within space, vividness, and a sense of nature [11]. In terms of living and architecture, elements such as traditional architectural style, family house style, architectural details, and the overall living environment were most often preferred by occupants [12–15]. With regard to the subjective characteristics of respondents, industry studies have demonstrated that factors that contribute significantly to and influence people’s visual aesthetic preferences for landscapes include occupation, education level, and gender [13].

With the gradual development of the new data environment, multidimensional exploration of the built environment at a finer scale across the country has become possible, advancing the need for more refined quantitative research on urban design. On the one hand, new analysis techniques are emerging, and the gradual application of GIS technology and Geodesign concepts in the field of urban design provides quantitative analysis tools, and spatial analysis supported by intelligent algorithms is expected to provide scientific support for spatial background identification and line selection effectiveness assessment. On the other hand, the emergence of high-precision open data such as road networks, points of interest, and building base maps provides new opportunities for the measurement of traditional unmeasurable data [16], and user profiles supported by multiple sources of data can also help to target accurate design. These new data and technologies can provide refined spatial feature extraction on the human scale, thus measuring various spatial quality that was previously difficult to assess accurately and offering new possibilities for human-centred urban design [17]. In this context, quantitative analysis of 5D elements, which were previously difficult to apply effectively in practice, has also emerged [18]. For example, location-based service (LBS) location data can be used in comparison to traditional measurement methods. Location-based service (LBS) location data can provide information on the behavioural activities of people over a wide range of time periods [19, 20]; street view images (SVIs) can be combined with machine learning; green visibility and street scape quality can be evaluated more efficiently and finely [21]; the combination of street base data and spatial design network analysis (sDNA) software enables large-scale and accurate street network accessibility measurements [22].

In summary, the rapid emergence of multisource urban data offers the possibility of in-depth analysis using the 5D concept to measure the quality of the built environment.

Most of the existing studies focus on a single dimension of built environment quality, while only a few researchers have combined two or more dimensions to measure street quality [23], and systematic analyses that integrate the five elements of the built environment at once are rare. In response to this research situation, this study aims to make use of new data and technologies to systematically measure many elements of street quality from a user-centred perspective and to compensate for the lack of scale and accuracy in traditional quantitative analysis.

2. Related Work

Although the scope of landscape visual evaluation has covered many aspects, most studies and cases in the past have been on large-scale natural landscapes; for example, some scholars have used example-validated methods such as psychological and psychological preferences to establish mathematical models for visual quantitative analysis and make a comprehensive evaluation of tourism resources from the perspective of visual aesthetics [24]. In recent years, the focus of foreign research on visual evaluation of landscapes has gradually shifted from the development of natural areas to the evaluation of urban and suburban areas, using visual evaluation methods to guide the planning and construction of cities. Especially after entering the 21st century, with the application of new technologies such as geographic information technology and computer image processing technology, the visual research and application of urban landscape, which was previously not easy to carry out due to the composition and influencing factors of urban landscape being too complex, has also developed greatly [25].

In the 1960s, some Japanese scholars explored the construction of greenery in the street space environment, and it was found that the green colour of plants can stimulate the human brain through vision; i.e., there is an obvious correlation between people’s evaluation of their feelings about the street and the street environment they are in [26]; through the study of the “Report on the Basic Survey of Greenery in Sugunami District,” it was confirmed that green vegetation can improve the satisfaction of the landscape in residential areas, and at the same time it also helps people to identify the landscape; countries such as Europe and the United States have also suggested that the amount of vegetation present in the visual field in large quantities contributes to the reporting of landscape evaluation. These research results are also gradually being recognised internationally, and a large number of scholars are beginning to experiment with and provide advice on the visual evaluation of landscapes in roads, parks, and residential areas, in order to improve the visual quality of landscapes in residential environments [27, 28].

In 1990, some scholars in China began to introduce theories related to visual evaluation from abroad, which laid a preliminary theoretical foundation for the development of visual landscape in China. In the same period, some scholars began to make preliminary explorations of examples in the evaluation of the visual environment of the landscape, for example, the exploration of landscape resource management

and visual impact assessment methods; from urban design theory, the basis was chosen to select impact factors to conduct research on the control of the visual environment of urban landscapes [29]. This also includes a large number of evaluations of the visual environment of landscapes relating to urban streets, such as the use of multifactor evaluation method to evaluate the visual landscape of roads; through multifactor evaluation method as well as questionnaire survey method to evaluate the visual landscape of urban roads; exploration of the applicability of urban road visual analysis; the use of landscape comprehensive evaluation index method to evaluate the quality of the landscape visual environment of Hong'an County Yingbin Avenue; the index of the visual evaluation of street landscape is divided into horizontal interface (street surface) and vertical interface (the three factors of street scape visual evaluation are divided into horizontal interface (street level), vertical interface (buildings, street trees), and street facilities); the visual index and SD method are proposed and applied to investigate urban streets, so as to obtain people's comprehensive evaluation of the visual environment of streets; two subjective evaluation methods, namely, the cognitive map method and the psychological evaluation method, are used to conduct field research on a number of living streets in Chongqing in the form of questionnaires, so as to derive the physical environment characteristics of street aesthetics.

In addition, modern information technology has also been widely used in the evaluation of the visual environment of the landscape at this stage in China, for example, the use of virtual reality technology and panoramic holographic technology to recreate the landscape along the Yellow Crane Tower in three dimensions, making the landscape simulation more in line with people's visual habits; the use of 3S technology to quantitatively evaluate the possible visual impact of new railways [30]; the use of Arc GIS visual field analysis function to analyse the visual field characteristics of the trails in the National Forest Park of Liaoning Province; the comprehensive evaluation of the aesthetic quality of Hangzhou Garden from the perspective of the public; the identification of green plant areas from Google Street View photos and the calculation of "green visual quality" at street scale. The study identifies areas of greenery from Google Street View photos, calculates a "green visual index" on the street scale, and suggests that street view photos could be of great use in future urban environmental planning. In general, the assessment methods and techniques used are constantly being developed and updated, and the objects and content of the assessment have been greatly expanded and deepened.

3. Study Site and Methodology

3.1. Study Site. The Shanghai area was chosen as the study site because local data was easily available and the area is known to be a multicultural gathering place where different cultures merge and are very prominent in reflecting the visual aesthetic preferences of different groups in the streetscape. With a planned area of 220 km^2 , the area is positioned as a regional leader in opening up to the outside

world and economic development, a base for high-tech industries, and a new modern urban area of high standard. Over the past 30 years, the urban space has developed from the north-east to the south-west, with the eastern part of the city being the birthplace of the development of the entire development zone [31]. The area of green space in the area has become more and more concentrated due to the needs of the development of the region, in contrast to the increasingly convenient transportation, the development of housing and public buildings, and the development of various industries.

3.2. Research Methodology. As mentioned earlier, one of the focuses of this study is to focus on the visual aesthetic preferences of fire-escapes at the subjective level. Therefore, subjective opinions from different groups of people had to be collected in order to determine local preferences for the visual aesthetics of the streets. Because opinions are open-ended and influenced by the personal preferences of the participants, this objective can best be achieved through the interview method of the qualitative research process [32]. And in landscape studies, extracting participants' perceptions through images has become a staple of the industry; this has proven to be an effective method. Building on this, Yohan Shao proposed that a subjective interview method using photographs as a basis is more effective in gathering opinions about the landscape [26]. Therefore, this study uses the experiential landscape annotation method in conjunction with the photo-oriented interview process and the basic principles of pinpoint notation to collect valid data for the study and analysis of the visual aesthetics of stroboscopes [28].

Firstly, the study area was set up along the main regional roads as reflected in the urban development map, with the route maintaining a good sense of spatial continuity, and the interviewing population would travel through the corresponding streets and record the characteristics of the visual experience of the street as perceived at each location. It is important to note that what is recorded are the more salient features at that location and not just the features that can be seen in the photographs, but also the psychological features etc. The photographs we use are simply a means of recording location and sequence (Figures 1 and 2).

These findings were then recorded in a table to clarify the quality of the different dimensions of visual aesthetics and the reasons for them. As the aim of this paper is to identify and assess the visual aesthetics of streets in urban areas, it was necessary for the participants to have some knowledge of the local area and therefore only local residents were selected for the data sampling. Local participants were selected from a mix of local construction and planning related professionals and nonprofessional local residents, a total of 90 people. It was also ensured that the participants covered three groups: older, middle-aged, and younger. The reason for this is that the different age groups have different insights into the local area and their interaction with the local environment, as shown in Figure 2.

In this study, the interviews were selected to be conducted at the regional site, in conjunction with a field visit



FIGURE 1: Selected photographic records of participants.



FIGURE 2: Photographs of some elements affecting visual perception.

survey. This method also proved to be interactive and effective in reducing the length of the interview, thereby reducing the resources required for the study [29]. Each set of interviews lasted between 40 min and 1 h, depending on the speed of the participant's choice and the details of the interview.

4. Study Design

4.1. Analytical Framework and Key Indicators. The study consisted of 4 main research steps: data collection, key feature extraction, alignment potential assessment, and feature mapping (Figure 3). Firstly, the study collected data from multiple sources such as LBS, POI, road network base data, MMT business hours data, SVIs, and refined built environment data for both sides of the Suzhou River area. Secondly, eight key variables representing the 5D theory were extracted from the dataset. In addition, as this is an implementation-oriented project, the maximum continuous nonmotorised section width of the road section is also used as a constructible indicator. Again, the relative weight of each of the nine factors in the six dimensions was calculated using AHP hierarchical analysis, and the indicators were then integrated to form an overall greenway suitability score. Finally, a radar map is used to draw a characteristic picture of a typical road section. The integrated multidimensional analysis can provide accurate guidance for the selection of greenway alignments and section design.

4.2. Study Case: Suzhou River Area, Shanghai. The study area covers the southern part of the Jing'an District on both sides of the Suzhou River, from North Henan Road in the east to Jiangning Road in the west, from West Beijing Road in the south to Jiaotong Road in the north, with a total land area of approximately 440 h m^2 . The Jing'an District is small in area and has a high population density of approximately $64,000 \text{ people/km}^2$, ranking first among all districts in Shanghai. However, at the end of 2016, the per capita parkland area in

Jing'an was $2.7 \text{ m}^2/\text{person}$, the third lowest in the central city, as shown in Figure 4. At the same time, the area has limited space for additional green space as there is less land available for construction [22]. Therefore, a study on the construction of urban greenways in this area can help to address the problems of environmental degradation and spatial disorder caused by high-density urban development and construction [23].

The activity density measurement was carried out based on Tencent's LBS location service data, and six representative time periods were selected on weekdays and weekends to capture the real-time crowd location distribution using Python computer language. A total of 20,405 data points were collected on weekdays and 15,436 data points were collected on weekends. As showed in Figure 5, each point represents a grid of 30 m in size, with high and low values representing the intensity of activity. The data were then preprocessed using ArcGIS for cleaning, spatial alignment, and kernel density analysis, and a spatial merging tool was used to translate the LBS into street segments with a 50 m buffer to reflect the current intensity of use in and around the street. This study shows that the buffer zone is a good reflection of the built environment characteristics of Shanghai.

For the street pedestrian count, a new method proposed by researchers from the City University of Hong Kong, Beijing Jiaotong University, and Tongji University was used, which showed good agreement with the field manual observation data in an empirical study in Tianjin (Figure 5). In this study, we firstly generated street scape sampling points with a sampling spacing of 30 m and calculated the latitude and longitude of each sampling point as well as the geographic orientation of the street segment where the sampling point was located; secondly, we set the street scape download parameters and obtained multiple phases of street scape for the sampling points in 2018–2019 by calling the APIs of Baidu and Tencent Street View [2]; then we cropped the street scape to cover only the pedestrian area on the left and right sides of the street as much as possible; finally, the LDCF algorithm is used to detect pedestrians. In addition, since the

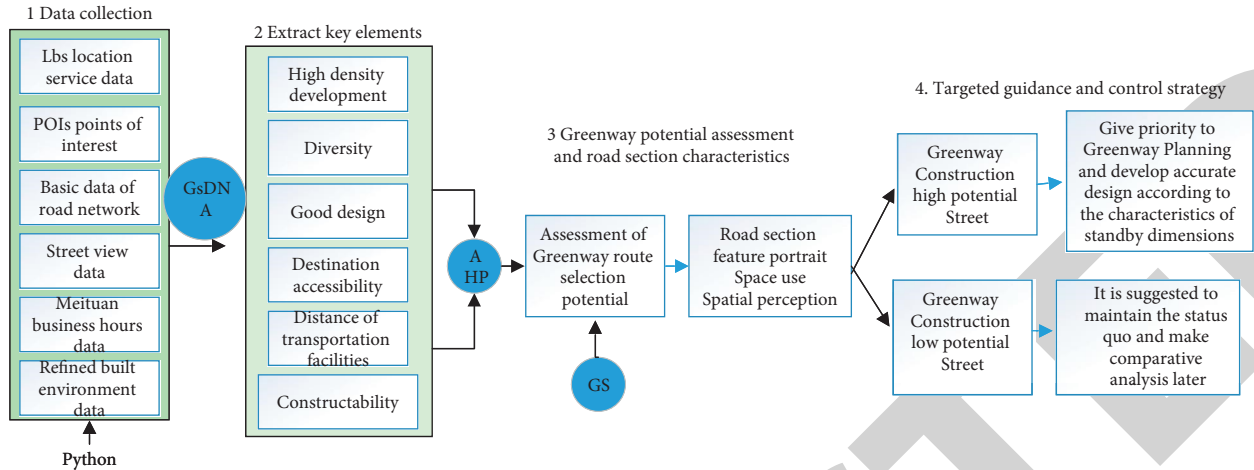


FIGURE 3: Research design framework.

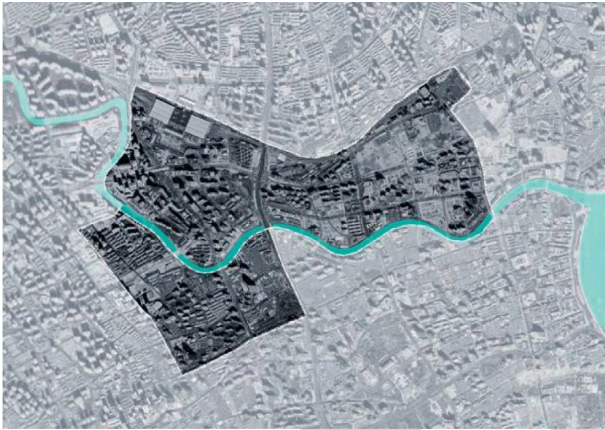


FIGURE 4: Scope of the study.

street view obtains the number of pedestrians at a certain point in time, and there are multiple sampling points in each street segment, and pedestrians are constantly moving in space and time, the average value of sampling points in each street segment is used instead of the pedestrian count in that street segment.

Diversity is measured in 2 ways, functional diversity based on POI data and average daily business hours based on Meituan data. With the popularity of mobile Internet, the increasing correlation between online data and physical facilities makes this data source an important resource for urban research. The study used Python and the Gaode Map API to extract 2,189 POIs in the study area, all of which were classified into different categories including food and beverage services, education, medical services, and transportation facilities, which can form a spatial distribution feature of various functional facilities (commercial, business, food and beverage, public services, etc.) in the city (Figure 6).

Once the data were obtained, functional diversity values were measured in 2 steps: (1) calculating the total number of POIs for each unit within a daily walking distance of 15 min and 1,000 m [25]; (2) measuring diversity using the widely

used entropy index [26], which has been applied in many built environment studies [13], calculated as

$$D_i = - \sum_{i=1}^R p_i \times \ln p_i, \quad (1)$$

where p_i is the proportion of the total number of POI types in a street unit, R is the total number of main functional categories, and the diversity value (D_i) is between 0 and 1. The greater the diversity index for each street unit, the better the mix of facilities. Average daily hours of operation based on Meituan data can reflect the economic vitality of the street. Since the hours of operation vary greatly from business to business, the hours of operation of shops are categorised as 06:00 am to 15:00 pm, 09:00–10:00 am to 15:00–18:00 pm, and 08:00–10:00 am to 18:00 pm: 00–10:00 am to 18:00–21:00 pm, 09:00–10:00 am to 21:00–22:00 pm, 06:00 am to 22:00 pm–0:00, 06:00 am to 0:00–03:00 am, 11:00–12:00 pm to 03:00–06:00 pm, 24 hours of operation in 8 categories (Figure 7). The shop hours were then aggregated to nearby streets to compare the overall hours of operation for each street.

5. Analysis of the Results

5.1. Results of the Analysis of Key Dimensions. For comparison purpose, the calculated results for the 9 subitems of the 6 dimensions related to urban greenway suitability were sorted into 5 categories from highest to lowest value, based on the ArcGIS quantile method. Figure 8 shows the scores for the eight variables from the five elements of the built environment in the study area, which measure the potential of each road section as an urban greenway alignment. In general, sections with higher development intensity and vibrancy are also more attractive to people and should be prioritised for inclusion in the slow-moving space system (Figure 8(a)). At the same time, streets with high functional diversity and longer operating hours offer a richer range of activities (Figure 8(b)). In addition, a higher quality of street space is physically and emotionally pleasing and can enhance the attractiveness of the street, making it more suitable

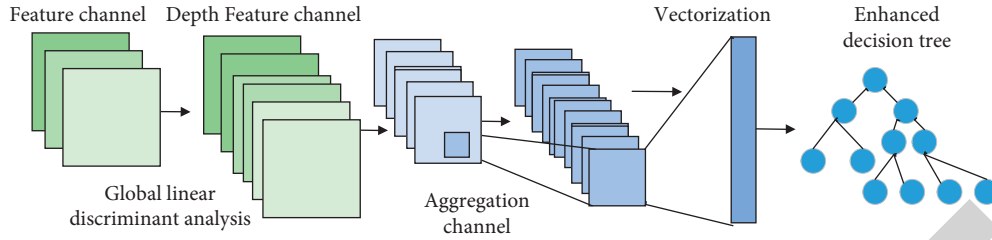


FIGURE 5: Example of street pedestrian count extraction.

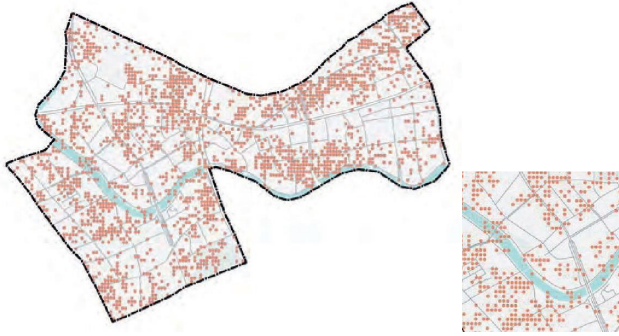


FIGURE 6: POI extraction example.



FIGURE 7: Example of street business hours distribution extraction.

for inclusion in the planned system of slow-moving space. There is no clear pattern in the intensity of development zones on the streets within the study area, and there are significant differences between road sections (Figure 8(c)). Considering the huge pulling effect of metro stations and the influence of pedestrian accessibility on pedestrian flow, priority should also be given to selecting streets with high scores on these two indicators when shaping the slow walking space system, which will help to improve the accessibility of the whole slow walking space system (Figure 8(d)).

Finally, the wider the nonmotorised road section width, the more likely it is to become a slow potential space under certain conditions. As shown in Figure 9, in the study area, the maximum continuous nonmotorised space is wider in sections such as Hengfeng Road, Hengtong Road, and Guangfu Road near the Hanzhong Road metro station on the north bank of Suzhou River, Qufu Road, North Xizang Road, Jinyuan Road, and Wuzhen Road near the Qufu Road

metro station, and Xinzha Road on the south bank of Suzhou River. Priority should be given to the inclusion of the slow-moving space system.

5.2. Weight Determination Based on Hierarchical Analysis.

By inviting more than 10 experts in the fields of architecture, planning, and landscape to conduct a comparison using the hierarchical analysis and evaluation method (AHP), the weight of the above-mentioned factors influencing the evaluation of the potential of urban greenway alignment was calculated. Based on the aforementioned analysis, nine factors in six dimensions were selected to build the indicator layer, and the importance of the indicators was compared by means of a two-by-two discriminant matrix, with the importance levels assigned on a scale of 1 to 9, with higher values indicating a higher importance of the former than the latter.

5.3. Greenway Alignment Potential Assessment.

Street data for a total of 9 indicators in 6 categories were further standardised so that the results were mapped to a range of values from 1 to 5, making the indicators comparable, and weighted and aggregated according to the weights of the elements determined by AHP above.

$$\text{Potential value} = \sum_{i=1}^9 D_i \times X_i, \quad (2)$$

where D_i are the influencing factors for each category of alignment potential from Table 1, and X_i are the corresponding weights.

On this basis, a comprehensive potential value for the urban greenway network can be calculated, the degree of which can be divided into five categories from high to low, arranged dark green to light green (Figure 10). The sections of Hengfeng Road, Hengtong Road, and Hanzhong Road near the Hanzhong Road Station on the north side of the Suzhou River, the Qipu Road Garment Wholesale Market at the Qufu Road Station, and the sections of Xinzha Road, Taixing Road, and Wuding Road on the west side of Wuding Road Street have high potential values for greenway construction, are coherent and have the potential to form vertical links with the Suzhou River, and have a higher degree of urgency for the construction of a slow-moving network, and should be given priority consideration for inclusion in the greenway planning system.

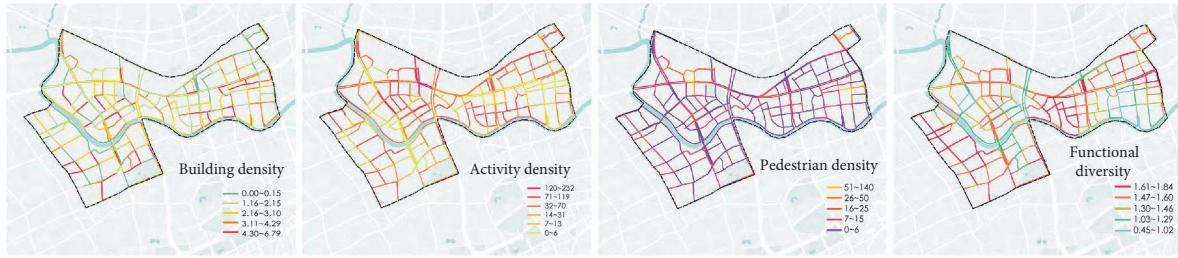


FIGURE 8: Distribution of the 8 categories of factors influencing the potential for greenway alignments.



FIGURE 9: Distribution of nonmotorised section widths.

TABLE 1: Assessment of urban greenway potential: impact weights for 6 dimensions and 9 subdimensions.

Key dimensions	Subitem	Weight (%)
Density	Activity density	10
	Pedestrian density	10
	Building density	5
Diversity	Functional diversity	15
	Business hours	20
Design	Visual quality	10
Distance between transportation facilities	Distance from subway	10
Destination accessibility	Section accessibility	10
Constructability	Continuous nonmotorised section width	10

5.4. *Characterisation of Road Sections with New Data and Technology.* This digital analysis model is used not only to assess the potential for route selection, but also to quickly generate a comprehensive profile of each section. It provides a scientific and systematic characterisation of the “how” of the greenway in a high-density built environment and allows for the subsequent design of each section to be targeted and to show the corresponding guiding strategies of human-scale characteristics.

Two of the more noteworthy categories in the aforementioned urban greenway potential assessment are presented here as case studies. The first category is the streets with the highest potential for greenway construction, such as Meiyuan Road (the section between Tianmu Xi Road and Republican

Road, Figure 11), which has taller buildings and high development intensity on both sides, a commercial type of building interface on both sides of the street, close proximity to the metro station Qufu Road Station, better spatial background conditions, narrower red lines, and nonmotorised widths, but wider pedestrian access, strong pedestrian accessibility, and suitability for slow walking. The spatial use vitality is high, with high activity, street vitality, and economic vitality, and a low distribution of flows in the current and planned context. Spatial perception of street quality is high, with a high level of green views, and street space enclosure is high, while pedestrian visibility is slightly lower.

The spatial context is good, with a high intensity of development and accessibility to public transport, but the



FIGURE 10: Distribution of potential values for the construction of urban greenways.

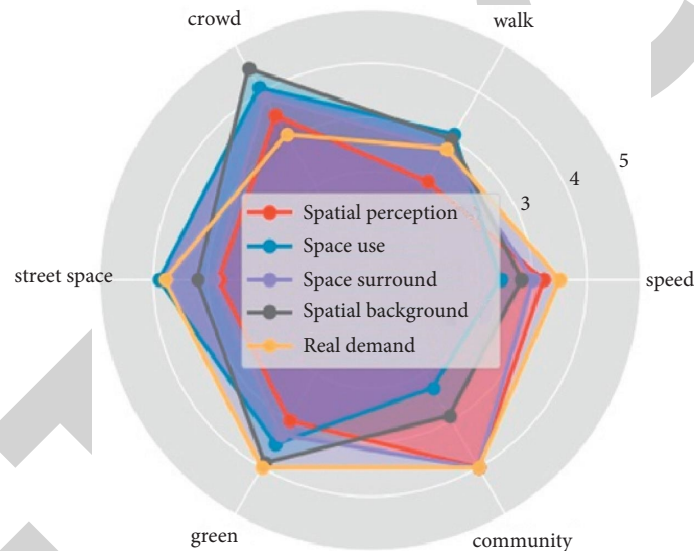


FIGURE 11: Characterisation of high potential sections of urban greenway construction.

proximity of green spaces and water bodies is poor and the accessibility to everyday life is poor. Overall, the spatial base is good, the spatial use is dynamic, the perceived quality of the space and the environment is excellent, and it only requires microrenewal at street level to improve its pedestrian accessibility and can be incorporated directly into the greenway planning system.

6. Conclusions

The analysis technique proposed in this study still has some limitations. Firstly, the current analysis is mainly applicable to the assessment of greenway alignment potential and road segment characteristics, but not for large-scale greenway selection, which needs to be explored in further research. Secondly, the reliability and validity of multiple Internet data, such as LBS and POI data, can be further improved. Future research can further integrate data from multiple

sources for calibration in order to improve the accuracy. In addition, the pretrained deep learning model used in this study is unable to distinguish specific types of green vegetation by the spatial features of plants in street scape images.

Data Availability

The dataset used in this paper is available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

- [1] M. Arsyad, I. B. Mochtar, N. E. Mochtar, and Y. F. Arifin, "Road embankment full-scale investigation on soft soil with

Research Article

Intelligent Detection and Analysis of Polycyclic Aromatic Hydrocarbons Based on Surface-Enhanced Raman Scattering Spectroscopy

Qian Zhang ¹, Bowen Chen,¹ Fazli Wahid,² Wanyun Feng,¹ and Xuerou Chen¹

¹College of Chemistry and Chemical Engineering, China West Normal University, Nanchong, Sichuan 637002, China

²Yangtze Delta Region Institute, University of Electronic Science and Technology of China, Huzhou, Zhejiang 313099, China

Correspondence should be addressed to Qian Zhang; zhangqian11@cwnu.edu.cn

Received 16 December 2021; Revised 9 January 2022; Accepted 15 January 2022; Published 2 February 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Qian 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.

Cycloaromatic hydrocarbons are a type of potentially hazardous chemicals that are widely present in the environment and pose a serious threat to human health. However, the traditional research methods for their detection process are cumbersome, the detection cycle is long, and the sensitivity is low. In response to the above problems, this article combines the molecular fingerprint information characteristics of surface-enhanced Raman scattering technology to simulate the four polycyclic aromatic hydrocarbons of pyrene, anthracene, phenanthrene, and trichene and quantitative detection of cyclic aromatic hydrocarbons and four kinds of polycyclic aromatic hydrocarbon mixtures. The experimental results show that the PAHs based on SERS have the advantages of higher sensitivity and high selectivity, which verifies the accuracy and feasibility of the method in this article.

1. Introduction

In recent years, more and more people have begun to pay attention to the importance of supramolecular interactions in nature. This type of role is mainly based on the use of artificially designed selective recognition substrates and target substances for molecular recognition [1]. The supramolecular system itself can be used in many fields such as drug sustained release and molecular sensing, and it can be used in combination with other technologies to take into account the applications of the two systems. Some researchers combine surface-enhanced Raman technology with some supramolecular bodies, such as calixarene [2] and cyclodextrin [3] for qualitative and quantitative detection and analysis. SERS technology has a more attractive aspect than Raman technology in terms of enhancement factor exceeding 10¹⁰. It is widely used in chemistry with its own high sensitivity [4], high selectivity [5], and no fluorescence [6] and the fields of biosensing and detection. Combining with supramolecular systems can make up for the inability of SERS technology to be applied to molecules that have no

physical or chemical interaction with the SERS substrate, inducing these molecules to combine with supramolecular systems to approach the area where the electromagnetic field is enhanced to determine the SERS signal of the molecule [2, 3].

Polycyclic aromatic hydrocarbons are organic compounds containing two or more benzene rings or heterocyclic rings and their substitutions. These compounds are hardly soluble in water and easily soluble in organic solvents. They are produced by incomplete combustion of organic polymer compounds such as coal, petroleum, wood, and tobacco. They are the main environmental and food pollutants, and a considerable part of them are carcinogenic [7]. Therefore, the detection of polycyclic aromatic hydrocarbons has always been a hot direction of scientific research. The existing PAHs detection technologies include gas chromatography [8, 9], high-performance liquid chromatography [10], chromatography spectrometry [11, 12], thin-layer chromatography, and capillary electrophoresis [13]. The advantages and disadvantages of immunoassays are shown in Table 1. Among them, the current mainstream

TABLE 1: Existing detection methods of PAHs and their advantages and disadvantages.

Detection method	Advantages	Disadvantages
Gas chromatography	High selectivity, high separation efficiency, high sensitivity, fast speed	Small application range, weak pertinence
High-performance liquid chromatography	Can analyze multiple compounds at the same time, good selectivity, high detection sensitivity	Solvent consumption is large, detectors have high separation efficiency, price is expensive
Chromatography and mass spectrometry	High separation efficiency, can determine the structural characteristics of the compound instrument	Expensive equipment, immature interface technology
Thin-layer chromatography	Simple, fast speed, intuitive	Low sensitivity, low separation efficiency
Capillary electrophoresis	High separation efficiency, simple operation	Sensitivity of existing detectors is not high
Immunoassay method	Strong pertinence, high sensitivity, low cost	Only one substance can be analyzed at a time, little information

detection methods for PAHs in the world are gas chromatography-mass spectrometry (GC-MS) and high-performance liquid chromatography (HPLC).

The characteristics and applicability of these detection methods are different, and there are many problems in the use process, such as long detection cycle, low sensitivity, and high cost. Therefore, in this paper, the SERS model combining cyclodextrin and gold nanoparticles is used for qualitative identification and trace detection of polycyclic aromatic hydrocarbons. The complex system of four polycyclic aromatic hydrocarbons (pyrene, anthracene, phenanthrene, and naphthalene) mixture is selected, and the design is used. The SERS-enhanced model is used for qualitative identification and quantitative detection. Experiments show that this method can be applied to the detection of PAHs mixture systems and has a certain application potential.

2. Related Work

Polycyclic aromatic hydrocarbons are a type of serious chemical pollutants. How to detect them quickly and accurately is an important scientific problem. In this paper, cyclodextrin is selected as the host molecule and modified on the surface of the SERS substrate. With the high sensitivity of SERS technology and the high selectivity of cyclodextrin, a quick and easy detection method for organic pollutants such as polycyclic aromatic hydrocarbons is developed, to achieve the purpose of qualitative identification and quantitative detection.

2.1. Surface-Enhanced Raman Spectroscopy. In the mid-1970s, the VanDuyne research group and the Creighton research group summarized experimentally and theoretically and found that this is a regular phenomenon based on rough surfaces, which is called surface-enhanced Raman scattering (surface-enhanced Raman scattering) (SERS) effect, that is, when the sample molecules are adsorbed (or approached) on the rough surface of metal materials such as gold, silver, and copper, the Raman signal can be greatly enhanced, and the method is established based on this enhanced effect with surface selectivity, it is called surface-enhanced Raman spectroscopy [14, 15]. With the continuous development of nanomaterial technology, transition metal has also achieved SERS response. After continuous development, surface-enhanced Raman spectroscopy can

realize single-molecule detection due to its high sensitivity; simple sample pretreatment, which can realize rapid on-site detection; good reproducibility, strong specificity, and advantages in environmental monitoring, food drug testing, biomedical analysis, etc. have been widely used [16, 17]. In surface-enhanced Raman spectroscopy, the Raman signal of the molecule to be measured is enhanced by exciting the local surface plasmon resonance (LSPR) of the substrate, so the magnitude of SERS enhancement mainly depends on the SERS substrate. However, the roughening of the substrate makes the preparation of the substrate a certain challenge. In the past few decades, in order to continuously improve the sensitivity, stability, and reproducibility of SERS detection, researchers have conducted a lot of research on SERS substrate materials [18, 19]. SERS technology requires the adsorption of analyte molecules onto the SERS substrate. After being adsorbed on the SERS substrate, the Raman signal of the analyte is enhanced. Unlike fluorescence, the spectral peaks obtained in SERS are narrower [20, 21].

The Raman shift characterizes the vibration characteristics of different groups in the molecule. Therefore, the Raman shift can be measured for qualitative and structural analysis of the molecule. Raman spectroscopy can be used for structural analysis of organic compounds. Due to different chemical environments, the Raman shifts of the same functional groups of different molecules have certain differences and will vary within a certain range, so it is convenient to distinguish various isomers, such as positional isomerism, geometric isomerism, cis-trans isomerism. For some groups, the Raman scattering signal is strong and the characteristics are obvious, which is also suitable for Raman determination. Raman spectroscopy can also be used for the study of polymer compounds, suitable for the determination of its geometric configuration, carbon chain skeleton or ring structure, and crystallinity. The polymer compound of the inorganic compound filler can be directly measured without separation. Raman spectroscopy is also an effective method for studying biological macromolecules, and it has been used to determine the structure of biochemical substances such as proteins, amino acids, carbohydrates, biological enzymes, and hormones. At the same time, Raman spectroscopy can study the composition, conformation, and interaction of biomolecules at extremely low concentrations close to the natural state. In addition, biological tissue sections such as eye lens, skin, and cancer tissue can be directly measured

without complicated processing. Therefore, Raman spectroscopy has been widely used in biology and medical research.

2.2. Detection Methods and Progress of Polycyclic Aromatic Hydrocarbons. At present, for the detection of PAHs in the environment and food, high performance liquid chromatography (HPLC) and gas chromatography-mass spectrometry (GC-MS) are mostly used, which are two more popular methods. High-performance liquid chromatography is a common separation technology. The method is used to determine polycyclic aromatic hydrocarbons. Generally, dichloromethane is used as the solvent, methanol water or acetonitrile-water is used as the mobile phase, and single or mixed solutions buffers, etc. The mobile phase is input into a chromatographic column equipped with a stationary phase. After the components in the column are successfully separated, it enters the fluorescence and ultraviolet detectors for detection to realize the analysis of the sample. This method has high separation efficiency, good selectivity, and high detection sensitivity. But the shortcomings are also particularly obvious. The solvent consumption is large, the types of detectors are few, the price is expensive, and the cost is high, and it can only detect organics with high boiling points and poor thermal stability. Zhang Qian et al. [22] constructed a method for detecting the residues of polycyclic aromatic hydrocarbons in the soil, using rapid solvent extraction, solid phase extraction, and gel permeation chromatography to purify the method, using high-performance liquid chromatography ultraviolet fluorescence. The detector is connected in series, and the method conforms to the national standard of measurement. He Yan [23] and others used high school liquid chromatography with ultraviolet and fluorescence detectors to determine 16 kinds of polycyclic aromatic hydrocarbons in the air, which is simple and quick to operate.

The advantage of gas chromatography-mass spectrometry (GC-MS) method is that the method can present good specificity and discrete monitoring ability in ion mode. GC-MS is different from high-performance liquid chromatography and is only suitable for analytes with poor volatility and good thermal stability. Gas chromatography has the advantages of high sensitivity and strong qualitative analysis reliability. Using gas as the mobile phase can achieve rapid equilibrium with the stationary phase, thereby achieving high precision and rapid analysis in a short time. Mass spectrometry, GC-MS has high resolution and lower detection limit for PAHs, which is recommended in many PAHs analysis methods. Liquid chromatography can realize the separation of nonvolatile substances and trace polar compounds. Although liquid chromatography cannot provide complete information of individual PAHs and alkyl-substituted PAHs like gas chromatography, it is coupled with fluorescence or mass spectrometry detectors. When used, the detection of a single PAHs in a complex mixture can be achieved. HPLC is a high-pressure, high-efficiency, high-speed, high sensitivity, and wide-ranging analysis technology. HPLC can significantly improve the sensitivity

of analysis by combining it with a fluorescence detector (FLD) or mass spectrometry detector. Due to the fluorescent properties of PAHs, HPLC is usually used in combination with FLD. ERS amplifies the Raman scattering signal of the target compound through the surface plasmon resonance phenomenon of metal nanoparticles and nanopatterned structures under the synergistic effect of electromagnetic and chemical effects. Due to the high hydrophobicity of PAHs, the affinity to the metal SERS active surface is low, resulting in low detection sensitivity of SERS. Functionalizing the surface of the substrate to increase the affinity for PAHs molecules is an important way to solve this problem. Capillary electrophoresis (capillary electrophoresis, CE) is an electrophoretic separation and analysis method that uses a capillary as a separation channel and a high electrostatic voltage field as the driving force. Capillary zone electrophoresis (CZE) is one of the most conventional CE methods. A CZE method based on an optimized cyclodextrin (CD) modification has been developed. This method is easier, faster, and more efficient. Be selective.

2.3. Application and Progress of Detection of Polycyclic Aromatic Hydrocarbons Based on Surface-Enhanced Raman Technology. At present, the detection methods for polycyclic aromatic hydrocarbons mainly include gas chromatography, liquid chromatography, high performance liquid chromatography and capillary electrophoresis. However, these methods require complicated pre-processing and time-consuming and labor-intensive, which brings difficulties to the general application of the method. Therefore, it is urgent to develop a fast, sensitive and simple detection method for qualitative identification and quantitative detection of PAHs. With the development of nano-preparation and characterization technology, SERS spectroscopy has made significant developments in both the basic principle research and the field of substrate preparation. Today, the sensitivity and reproducibility of many new SERS substrates have reached practical standards. However, for those molecules that do not have substituent groups that can interact with precious metal substrates, it is often necessary to use other modification methods to modify the SERS substrate to perform SERS detection on these molecules. For the SERS detection of PAHs, the methods of modifying the SERS substrate can be roughly divided into five categories:

Long-chain alkanes modified SERS substrates SERS humic acid modified SERS substrates; calixarene modified SERS substrates; amethyst dication-modified SERS substrates and mercapto-substituted cyclodextrin-modified SERS substrates. Regardless of the modification method of the SERS substrate, the general idea is to use the interaction between the PAHs molecules and the surface modification molecules of the SERS substrate to bring them close to the surface of the modified substrate, and then reach the enhanced area of the enhanced substrate surface, so as to complete the modification. SERS detection. Several SERS substrate modification methods will be introduced and compared below. The specific PAH molecules detected by several modified substrates and the corresponding detection limits are listed in Table 2.

TABLE 2: Detection limit corresponding to polycyclic aromatic hydrocarbon molecules.

Modification materials	Detection molecule	Detection limit/M
Long chain alkanes	Pyrene	10^{-8}
	Naphthalene	10^{-7}
	Phenanthrene	10^{-7}
Viologendication	Pyrene	10^{-9}
	Pyrene	10^{-8}
Calixarene	Triphenylene	10^{-9}
	Coronene	10^{-10}
	Anthracene	10^{-7}
Cyclodextrin	Pyrene	10^{-8}
	Triphenylene	10^{-6}

Surface-enhanced Raman spectroscopy still has some problems in the detection of PAHs based on SERS substrates modified by supramolecular chemistry. The long-chain alkane-modified SERS substrates use alkyl substituents to modify the noble metal SERS substrates, mainly using the hydrophobic interaction between PAHs and long-chain alkanes. This method can obtain high detection sensitivity, but the detection selectivity is poor. Humic acid is a macromolecular organic acid composed of aromatics and various functional groups, which has good physiological activity and functions such as absorption, complexation, and exchange. The detection of PAHs on the humic acid-modified SERS substrate is also based on the hydrophobic interaction between the two, so it also encounters the problem of poor selectivity. The interaction between the viologen dication-modified SERS substrate and PAHs, in addition to the hydrophobic effect, also includes the π - π interaction between the aromatic rings. This method helps to improve the adsorption selectivity of the SERS substrate to the measured substance. The use of supramolecular chemistry to modify the surface of SERS substrates is an important development in the detection of PAHs. Future research is expected to focus on two aspects. On the one hand, further improve the reproducibility of the SERS substrate, combined with stoichiometric methods. On the other hand, it can be combined with other analytical techniques, especially separation techniques, to make the qualitative and quantitative analysis of SERS spectroscopy more perfect.

3. Experiment and Analysis

The 60 nm gold sol was purchased from BB International with a concentration of 3.4×10^{-11} M. Chemical reagents such as anthracene, pyrene, phenanthrene, and pyrene were purchased from Wako, Japan. The experimental water is tertiary deionized water. First, the gold sol was mixed with sulfhydryl-substituted cyclodextrins of different concentrations overnight, and then the ethanol stock solution of polycyclic aromatic hydrocarbons was diluted with water to the desired concentration and mixed overnight. Finally, the polycyclic aromatic hydrocarbons entering the cyclodextrin cavity are centrifuged with different centrifugal speeds.

3.1. Experimental Instrument. The Raman spectrometers used in the experiment were all from Photon Design, and the 488 nm argon ion laser was used as the excitation light source. The power of the laser finally reaching the sample is about 1 mW. All Raman and spectra adopt Rubberband method for baseline correction.

Pyrene is a light yellow (pure product is colorless) solid aromatic compound. Its molecular formula is $C_{16}H_{10}$. It can also be oxidized to obtain 1,4,5,8-naphthalene tetracarboxylic acid is used to make fuels, synthetic resins, and engineering plastics; in addition, it can be acylated with > 1 winter pyrene to obtain vat dyes.

Anthracene is a light yellow needle-like crystal with light blue fluorescence (pure band purple fluorescence), so it has another name “flashing crystal”, and its molecular formula is $C_{14}H_{10}$.

Phenanthrene is a white flaky crystal whose molecular formula is $C_{14}H_{10}$. Phenanthrene and anthracene are isomers, so they have many similarities with anthracene.

Naphthalene is a colorless flake crystal with a special smell, and its molecular formula is $C_{10}H_8$. Naphthalene has the characteristics of being volatile and easy to sublime, so it must be kept in a sealed container.

3.2. Experimental Results and Analysis. In this experiment, gold nanoparticles modified with sulfhydryl-substituted cyclodextrin were used to detect polycyclic aromatic hydrocarbons. In order to judge the influence of the coverage of the sulfhydryl-substituted cyclodextrin on the surface of the gold nanoparticles on the detection of polycyclic aromatic hydrocarbons, we selected different concentrations of supramolecular bodies from 10^{-4} to 10^{-9} M and mixed with the same amount of gold sol. Then 10-5M anthracene molecule is used as the detection molecule, and SERS detection is performed on it. The peak intensity at 1398 cm^{-1} is selected as the research object, and the curve of the peak intensity at 1398 cm^{-1} is drawn with the change of the concentration of sulfhydryl-substituted cyclodextrin. The result is shown in Figure 1. As the concentration of supramolecular host increases, the intensity of the SERS peak of anthracene first increases and then decreases. This phenomenon is related to the optimal coverage of sulfhydryl-substituted cyclodextrin on the surface of gold nanoparticles. The experimental result is that the best SERS enhancement effect is achieved when the concentration of cyclodextrin is 10^{-6} M.

In order to theoretically verify the above experimental results, we have carried out theoretical calculations on the surface coverage of the system. First, we assume that the best SERS enhancement effect is achieved when the gold nanoparticle surface is completely covered. From formulas (1)–(3) and the size of the gold nanoparticles and the size of the cyclodextrin, we can roughly calculate that the concentration when the thiol-substituted cyclodextrin completely covers the gold nanoparticles is 12×10^{-6} M. The theoretical calculation results are in good agreement with the experiment.

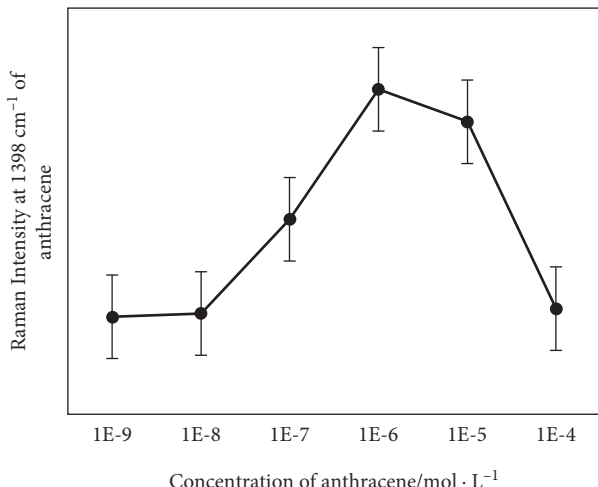


FIGURE 1: SERS intensity of an anthracene marker band at 1398 cm^{-1} versus the concentration of CD-SH.

$$S_{\text{gold}} = S_{S-\beta\text{CD}} = 4\pi r^2, \quad (1)$$

$$\frac{S_{\text{gold}}}{S_{S-\beta\text{CD}}} = \frac{C_{S-\beta\text{CD}}}{C_{\text{gold}}}, \quad (2)$$

$$C_{S-\beta\text{CD}} = \left(\frac{S_{\text{gold}}}{S_{S-\beta\text{CD}}} \right) * C_{\text{gold}}. \quad (3)$$

Among them, S_{gold} and $S_{S-\beta\text{CD}}$ are the surface areas of gold nanoparticles and sulfhydryl-substituted cyclodextrin, respectively, and C_{gold} and $C_{S-\beta\text{CD}}$ are the molar concentrations of gold nanoparticles and sulfhydryl-substituted cyclodextrin, respectively. When the concentration of sulfhydryl-substituted cyclodextrin was lower than $1 \times 10^{-6}\text{M}$, the SERS signal of PAHs increased as the amount of cyclodextrin adsorbed on the surface of gold nanoparticles increased. This is because more mercapto-substituted cyclodextrin means that more cyclodextrin PAHs host-guests are adsorbed on gold nanoparticles, and more PAHs enter the enhanced area of gold nanoparticles. In turn, a stronger SERS signal can be obtained. The best effect is when the concentration of sulfhydryl-substituted cyclodextrin is equal to $1 \times 10^{-6}\text{M}$. When the concentration of sulfhydryl-substituted cyclodextrin is greater than $1 \times 10^{-6}\text{M}$, these molecules are combined with PAHs, so that excess cyclodextrin molecules and PAH molecules will be removed during centrifugation. As a result, the strength is reduced.

4. SERS Qualitative Identification and Quantitative Detection of PAHs

In order to achieve the best results in the detection of PAHs, we used unequal centrifugal speeds to enrich PAHs, and then select optimum centrifugal speed. The 10^{-5}M anthracene molecule was used as the detection molecule, and other experimental conditions were unchanged. The centrifugal speed was from 1000 to 2500 rpm/s. After centrifugation, the sediment was retained to remove the supernatant, and the operation was repeated twice after washing with water.

4.1. Quantitative Analysis Components of PAHs by SERS Spectroscopy. We have done comparative experiments in the previous section to prove that this SERS signal is indeed derived from the supramolecular modified SERS substrate we designed instead of the precious metal itself. For the four types of PAHs we studied, the measured SERS signals are in good agreement with each solid. Figure 2, respectively, represents the dependence of the peak intensity of several characteristic peaks selected from each SERS spectrum of the four PAHs, anthracene, pyrene, phenanthrene, and phenanthrene, as a function of concentration. It can be seen from the figure that the SERS intensity of the four polycyclic aromatic hydrocarbons shows almost the same trend as the concentration decreases. In addition, the detection limits for these four polycyclic aromatics are 100, 10, 100, and 1000 nM. Using this modification method, the detection limits for anthracene and pyrene molecules are increased by 100 times.

From Figure 2, we can intuitively compare the detection capabilities of the four polycyclic aromatic hydrocarbons. The relationship between the detection limits is pyrene > anthracene > triphthone > phenanthrene. This can be explained by the size matching effect between polycyclic aromatic hydrocarbons and the inner cavity of cyclodextrin. The inner cavity of cyclodextrin has the largest molecular size of benzene. It is difficult to enter the inner cavity of cyclodextrin. In addition, after centrifugation, there is almost no signal from the benzene molecule. For the pyrene molecule, it is likely to form a composite structure of subject-object-subject. This composite structure uses the pyrene molecule as a “molecular bridge” to connect two gold particles, which forms the “hot spot effect” of SERS to enhance the signal of the pyrene molecule. It can be inferred from this that the SERS effect obtained by using the SERS model is closely related to the size matching effect of the polycyclic aromatic hydrocarbon molecules and the cyclodextrin cavity.

4.2. SERS Qualitative Identification and Quantitative Detection of PAHs Mixture. In this paper, sulfhydryl-substituted cyclodextrin-modified gold substrates have been used to obtain SERS spectra with better signal to noise ratios for SERS detection of single polycyclic aromatic hydrocarbons. Next, continue to study the qualitative identification of the mixture of four PAHs. We can clearly distinguish the existence of each PAHs, and there are some new SERS peaks in the figure. On the one hand, we infer that these are due to the four PAHs, each of which competes to enter the cyclodextrin cavity, which makes the spectrum more complicated. On the other hand, when the concentration of PAHs decreases, the noise is relatively large, and some impurity peaks are more obvious. We fixed the concentration of the three polycyclic aromatic hydrocarbons to remain unchanged, changed the content of one of the polycyclic aromatic hydrocarbons, and used the cyclodextrin-modified SERS substrate designed in this experiment to perform SERS quantitative detection of the mixture system. We changed the concentrations of anthracene and pyrene molecules from 500 to $50\mu\text{M}$ to

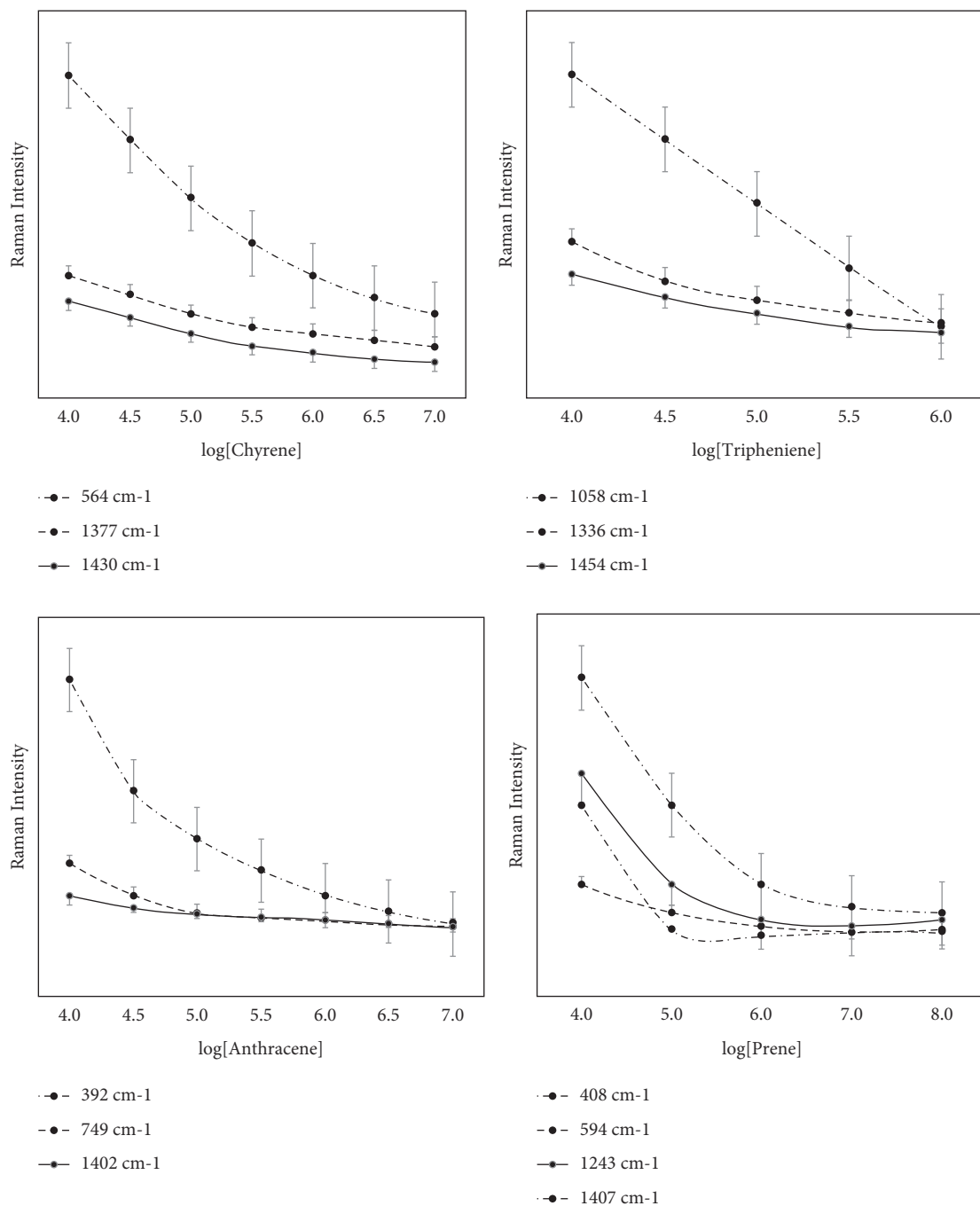


FIGURE 2: SERS intensity at discriminant peaks versus concentrations for anthracene, pyrene, chrysene, and triphenylene with CD-SH functionalized AuNP.

maintain the concentration of the other three polycyclic aromatic hydrocarbons at $10\mu\text{M}$ and extracted the peak intensity at 1539cm^{-1} , 1259cm^{-1} , and 390cm^{-1} as a function of concentration. The intensity of each peak of anthracene and pyrene in the mixture exhibits roughly the same law as the concentration changes, which proves that this method can be applied to the detection of polycyclic aromatic hydrocarbon mixture systems and has a certain application potential.

5. Conclusions

In this article, we designed a SERS substrate that combines supramolecular action, that is, a gold substrate modified by sulfhydryl-substituted cyclodextrin. This substrate has the characteristics of stability and long storage time, making it similar to polycyclic aromatic hydrocarbons that have no effect on gold. Of molecules are detected using SERS technology. The effect of the surface coverage of sulfhydryl-

substituted cyclodextrin on the effect of gold nanoparticle surface on SERS and the effect of centrifugal speed on the results of SERS are discussed in detail. It has realized not only the quantitative detection of single components of pyrene, anthracene, phenanthrene, and benzophenone but also the quantitative detection of a variety of complex systems of polycyclic aromatic hydrocarbons. The experimental results show that the detection of PAHs based on SERS has the advantages of higher sensitivity and high selectivity, which verifies the accuracy and feasibility of this method. SERS-based PAH detection methods have promising applications in analytical chemistry and environmental science research, but these are mostly limited to basic and laboratory research. The real application of SERS in practical environmental science awaits further development of large-area-ordered SERS active substrates and further communication and cooperation between chemists and environmental scientists in this field.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] I. Beletskaya, V. S. Tyurin, A. Y. Tsivadze, R. Guilard, and C. Stern, "Supramolecular chemistry of m," *Chemical Reviews*, vol. 109, no. 5, pp. 1659–1713, 2009.
- [2] L. Guerrini, J. V. Garcia-Ramos, C. Domingo, and S. Sanchez-Cortes, "Sensing polycyclic aromatic hydrocarbons with d-functionalized Ag nanoparticles by surface-enhanced Raman scattering," *Analytical Chemistry*, vol. 81, no. 3, pp. 953–960, 2009.
- [3] Y. Xie, X. Wang, X. Han et al., "Sensing of polycyclic aromatic hydrocarbons with cyclodextrin inclusion complexes on silver nanoparticles by surface-enhanced Raman scattering," *The Analyst*, vol. 135, no. 6, pp. 1389–1394, 2010.
- [4] X. X. Han, B. Zhao, and Y. Ozaki, "Surface-enhanced Raman scattering for protein detection," *Analytical and Bioanalytical Chemistry*, vol. 394, no. 7, pp. 1719–1727, 2009.
- [5] J. A. Dieringer, A. D. McFarland, and N. C. Shah, "Surface-enhanced resonance Raman scattering revealed by an analysis of anti-stokes and stokes Raman spectra," *Physical Review B*, vol. 76, no. 12, pp. 76–85, 2007.
- [6] G. G. Huang, M. K. Hossain, X. X. Han, and Y. Ozaki, "A novel reversed reporting agent method for surface-enhanced Raman scattering; highly sensitive detection of glutathione in aqueous solutions," *The Analyst*, vol. 134, no. 12, pp. 2468–2474, 2009.
- [7] Q. W. Xuan, "The methods and progress of polycyclic aromatic hydrocarbons detection," *Chinese Journal of Environmental Management*, no. 3, pp. 49–52, 2013.
- [8] Q. Miao, W. Kong, S. Yang, and M. Yang, "Rapid analysis of multi-pesticide residues in lotus seeds by a modified QuEChERS-based extraction and GC-ECD," *Chemosphere*, vol. 91, no. 7, pp. 955–962, 2013.
- [9] C. I. Ng and X. Zhang, "A quick analytical method using direct solid sample introduction and GC-ECD for pesticide residues analysis in crops," *Talanta*, vol. 85, no. 4, pp. 1766–1771, 2011.
- [10] D. Harshit, K. Charmy, and P. Nrupesh, "Organophosphorus pesticides determination by novel HPLC and spectrophotometric method," *Food Chemistry*, vol. 230, pp. 448–453, 2017.
- [11] M. V. Barbieri, C. Postigo, N. Guillem-argiles et al., "Analysis of 52 pesticides in fresh fish muscle by QuEChERS extraction followed by LC-MS/MS determination," *The Science of the Total Environment*, vol. 653, pp. 958–967, 2019.
- [12] B. Y. Durak, D. S. Chormey, M. Firat, and S. Bakirdere, "Validation of ultrasonic-assisted switchable solvent liquid phase microextraction for trace determination of hormones and organochlorine pesticides by gc-ms and combination with quechers," *Food Chemistry*, vol. 305, 2020.
- [13] J. Sun, C. Zheng, X. Xiao, L. Niu, T. You, and E. Wang, "Electrochemical detection of methimazole by capillary zone electrophoresis at a carbon fiber microdisk electrode," *Electroanalysis*, vol. 17, pp. 1675–1680, 2010.
- [14] Z.-Q. Tian, B. Ren, and D.-Y. Wu, "Surface-enhanced Raman scattering: from noble to transition metals and from rough surfaces to ordered nanostructures," *The Journal of Physical Chemistry B*, vol. 106, no. 37, pp. 9463–9483, 2002.
- [15] C. Ana Paula, S. Franca, and J. Irudaráj, "Surface-enhanced Raman spectroscopy applied to food safety," *Annual Review Food Science Technology*, vol. 4, 2013.
- [16] R. Botta, G. Upender, R. Sathyavathi, D. Narayana Rao, and C. Bansal, "Silver nanoclusters films for single molecule detection using surface enhanced raman scattering (SERS)," *Materials Chemistry and Physics*, vol. 137, no. 3, pp. 699–703, 2013.
- [17] R. Botta, Y. Liu, H. Zhou et al., "Label and label-free based surface-enhanced Raman scattering for pathogen bacteria detection: a review," *Biosensors and Bioelectronics*, vol. 94, pp. 131–140, 2017.
- [18] Z. Han, S. Lin, Z. Yuan et al., "Rapid determination of illegal additives chrysoidin and malachite green by surface-enhanced raman scattering with silanized support based substrate," *Chinese Chemical Letters*, vol. 29, 2018.
- [19] J. R. Lombardi and R. L. Birke, "A unified approach to surface-enhanced Raman spectroscopy," *Journal of Physical Chemistry C*, vol. 112, no. 14, pp. 5605–5617, 2008.
- [20] G. Mcnay, D. Eustace, W. E. Smith, K. Faulds, and D. Graham, "Surface-enhanced Raman scattering (sers) and surface-enhanced resonance Raman scattering," *Applied Spectroscopy*, vol. 65, 2011.
- [21] M. D. Porter, R. J. Lipert, L. M. Siperko, G. Wang, and R. Narayanan, "SERS as a bioassay platform: fundamentals, design, and applications," *Chemical Society Reviews*, vol. 37, no. 5, pp. 1001–1011, 2008.
- [22] Z. Qian, "Determination of polycyclic aromatic hydrocarbons in soil by ase-gpc-spe purification and HPLC detection," *Environmental Chemistry*, vol. 4, pp. 771–777, 2011.
- [23] H. Yan, "HPLC determination of 16 different polycyclic aromatic hydrocarbons in air," *Modern Preventive Medicine*, vol. 42, no. 8, pp. 1475–1500, 2015.

Research Article

Research on MOOC Teaching Mode in Higher Education Based on Deep Learning

Yuan Tian,¹ Yingjie Sun,¹ Lijing Zhang,¹ and Wanqiang Qi ²

¹Aviation University of Air Force, Changchun, Jilin 130022, China

²School of Automotive Engineering, Jilin Teachers Institute of Engineering and Technology, Changchun, Jilin 130022, China

Correspondence should be addressed to Wanqiang Qi; qwqwall@jlenu.edu.cn

Received 29 November 2021; Revised 15 December 2021; Accepted 23 December 2021; Published 29 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yuan Tian 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 development of computer technology and network technology and the widespread popularity of electronic equipment, communication among people is more dependent on the Internet. The Internet has brought great convenience to people's lives and work, and the Internet data is constantly being recorded. People's data information and behavior information, which provides the basis for data mining and recommendation systems, mining users' information and behaviors, and providing "user portraits" for each user, can provide better services to users and it is also an important part of the recommendation system. In one step, this article takes MOOC education resources as the research goal. In order to improve the effective management of MOOC platform resources based on traditional methods, this article combines relevant data sets and recommendation techniques to initially build a learning platform, implements a deep neural network algorithm, and recommends related services. The request and response data were explained, and through the online learning data set, based on the learner's historical learning records, the learning resources were simulated and recommended to the learners. The resource customization module was elaborated. Through the results of resource recommendation, a personalized learning resource recommendation platform was initially realized, which more intuitively demonstrated the recommendation effect and better realized the teaching management of the MOOC platform.

1. Introduction

MOOC stands for "Massive Open Online Course," which is an open-access and large-scale participation online course. This brand-new course model was opened by Professor David Wiley of Utah State University in the United States and Professor Alec Couros of the University of Regina in Canada. The two Internet open courses developed and evolved. In recent years, "Internet+" education has been surging. With the advantages of breaking time and space constraints and integrating and optimizing learning resources, MOOC has gradually gained widespread attention and developed rapidly in China. MOOC has changed the traditional classroom structure where teachers mainly teach and changed the single mode of acquiring new knowledge through books, making the identities of teachers and students fundamentally change in the teaching process.

The most popular teaching reform in today's education field is the rise of large-scale open online course MOOC platforms. We have investigated the types of courses offered on the existing MOOC platforms in China and found common problems. The course resources of the major platforms are not balanced. They mainly offer engineering theoretical courses but lack experimental courses corresponding to engineering theoretical courses. According to the statistics of the "Chinese University Online Open Course Forum," there are currently more than 1,300 online courses on the "China University MOOC" platform, the largest open online course platform in China, and more than 20 million courses have been selected. However, according to the syllabus of engineering disciplines, there are more than 360 courses that need to be opened, but only 11 results of online experimental courses were actually searched, which only accounted for 3.05% of the total

number of experimental courses specified in the syllabus. The proportion of experimental courses offered by various platforms is seriously unbalanced [1–5].

With the rapid development of the MOOC platform, more and more college students and social learners choose courses, and college teaching also attaches great importance to the design of online courses. Especially for engineering courses, the courses of some colleges and universities have supported teaching links such as credit certification, homework release, course performance evaluation, and academic certificate issuance. The final subject scores are evaluated based on the online teaching process, which completely replaces the traditional classroom teaching model. Judging from this development trend, it also fully illustrates the importance of course content design under the online platform. However, online learning led by MOOC still faces many doubts. Among them, the most criticized is that there are not many learners who insist on completing the courses, and the dropout rate remains high. Online learning does have a huge impact on traditional teaching with its unique advantages, but we also need to note that the nature of online learning determines that teachers and students are separated in time and space, and teachers cannot observe and learn as intuitively as in traditional classrooms. As a result, the learning behavior of the learner cannot detect and feed back problems in time and give early warning. Therefore, based on the theory of social cognition, this research focuses on the learning environment and learning psychology of the two MOOC learning context elements and constructs a MOOC learning behavior model oriented by the cognitive participation of learning behavior. The relationship between the effects is analyzed in depth, and the teaching rules and pedagogical principles are explained. Finally, based on the model, the deep neural factorization machine is used to predict the learning effects of MOOC learners.

2. Related Theoretical Methods

2.1. Online Learning Behavior. Compared with the learning behaviors generated in traditional classrooms, MOOC online learning behaviors have their own outstanding characteristics. They are not restricted by space, learning resources are richer, and learning is more intelligent and efficient. During the interaction between learners and the Internet, the behaviors generated are easy to collect, mine, and analyze, which can provide data and theoretical basis for the improvement of the level of online education in the future.

The effective learning behaviors of MOOC learners mainly include the following:

- (1) *Watch the Course Video.* Watching the course video is the main means for learners to learn in MOOC. Through the learning of video resources, learners can follow the teacher to intuitively learn the content of the course.
- (2) *Submit Homework and Test.* In the learning platform, as the course progresses, teachers will assign some

homework for learners to answer. The homework can not only check the learner's learning effect in time, but also broaden the learner's horizons, so that the learner can summarize and reflect. With the increase in learning content, teachers hope to understand learners' learning effects through tests and examinations and at the same time, through tests and examinations, to give learners the opportunity to check for deficiencies.

- (3) *Participate in Exchanges in the Discussion Forum.* After the learners have finished learning the video of the course, they may still have doubts about some of the problems of the course. At this time, they can go to the discussion area corresponding to the course to search for the problem and see if any other learners have encountered the same problem. You can also post by yourself, seek help from other learners, or answer questions raised by other learners in the forum to help everyone make progress together [6–10].
- (4) *Other Learning Behaviors.* When the learner is studying the MOOC course, in addition to the learning of video resources, the teacher will also provide learners with other forms of learning resources, such as courseware, bibliography, sample codes, etc., which can be used as a supplement to video learning, which is very important learning behavior.

In order to analyze and predict learners more accurately, 11 learning behavior characteristics that can represent their learning behavior characteristics are selected from the behavior logs generated by the learners in the MOOC learning process. These 11 learning characteristics are selected from the number of activities. Look at the video situation, homework situation, forum discussion situation, and other dimensions to describe the learner's learning behavior. The names and descriptions of the 11 learning behavior characteristics are shown in Table 1.

2.2. Text Classification on MOOC Platform. This article mainly uses LSTM in deep learning to classify texts in MOOC platform teaching management and innovates on the basis of RNN neural network model. The biggest difference from other models is that the biggest feature of recurrent neural network is that it propagates in the forward direction. In the process, the data obtained by the current neuron is affected not only by the current input data but also by the output data of the previous node. Based on this feature, the cyclic neural network can better process data with time series attributes. The standard RNN network structure is shown in Figure 1, where X represents the input, O represents the network output, U , V , and W , respectively, represent the weights between the layers, and S represents the output of the hidden layer in the network [11–13].

Numerous experiments have shown that RNN has certain shortcomings for dealing with long-time sequence problems. The longer the time series is, the weaker the

TABLE 1: The descriptions of the 11 learning behavior characteristics.

Label	Name	Illustration
D ₁	Number of courses visited	Number of visits to the course during the MOOC study
D ₂	Duration of the course	Duration of access to courses during MOOC study
D ₃	Access to other modules of the course	Number of visits to other modules of the course during the MOOC study
D ₄	Number of visits to the forum	Number of visits to assignments during MOOC learning
D ₅	Visits to the wiki	The number of visits to the forum corresponding to the course in the MOOC study course
D ₆	Total number of visits	Number of visits to Wikipedia during MOOC study
D ₇	Activities	Active days in the MOOC learning process
D ₈	Active days	The total number of learning activities in the MOOC learning process
D ₉	Page closed	The number of times the web page was closed during the MOOC learning process
D ₁₀	Video views	The number of times the instructional video was watched during the MOOC learning process
D ₁₁	Watch video time	The number of times the web page was closed during the MOOC learning process

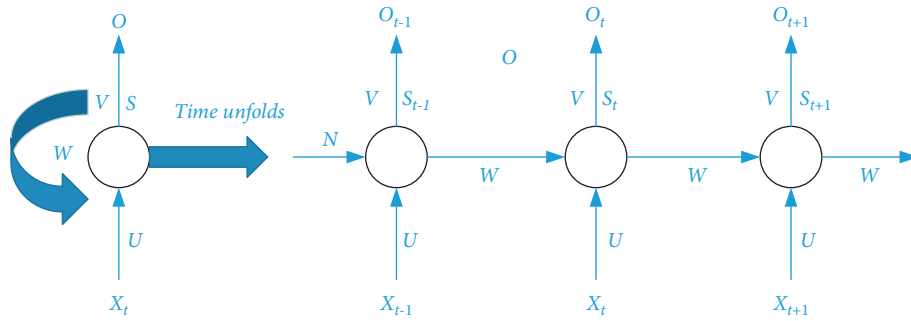


FIGURE 1: Standard structure diagram of cyclic neural network.

memory capacity of the RNN will become, which will cause the network to be unable to save the previous output information when subsequent neurons are trained. In order to solve such problems, the researchers put forward an LSTM model that can record long-term information through experiments and demonstrations. Early methods to solve these problems were either tailored to specific problems or did not extend to long-term dependencies. Only LSTM is both versatile and effective in capturing long-term time dependencies. LSTM implements the screening of information through the “gate” control structure. The “gate” is represented as a fully connected layer in the network. The standard LSTM core unit consists of an input gate, an output gate, and a forget gate. With 3 kinds of gate control composition, each gate has a different function, the input gate is mainly responsible for determining how much valid input information is retained, the output gate determines how many valid unit states are output at the current moment, and the forget gate determines whether to retain the previous one. The specific network structure of the status output information is shown in Figure 2 [14].

2.3. Evaluation Index of the Algorithm. TOP-N recommendation is to recommend the Top-N items that users of the MOOC platform may like. Generally, accuracy and recall rates are commonly used to measure the effect of Top-N recommendation. The specific definition and expression are as follows:

$R(u)$ represents the recommendation list obtained by the target user on the training set, and $T(u)$ represents the recommendation list obtained by the user on the test set. According to the above information, the recall rate is expressed as

$$\text{recall} = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|}. \quad (1)$$

The accuracy of the recommended results is expressed as

$$\text{precision} = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|}. \quad (2)$$

Novelty is defined as the reciprocal of popularity. If a product is very popular, it appears more frequently, which is not a novel recommendation result for users. It is expressed as formula (3), where K_c represents the popularity of the product.

$$\text{novelty} = \frac{1}{\log_2(K_c + 1) + 1}. \quad (3)$$

In the Top-N recommendation, the F1 comprehensive evaluation index is also the ratio of the precision rate and the recall rate, expressed as

$$F1 = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}. \quad (4)$$

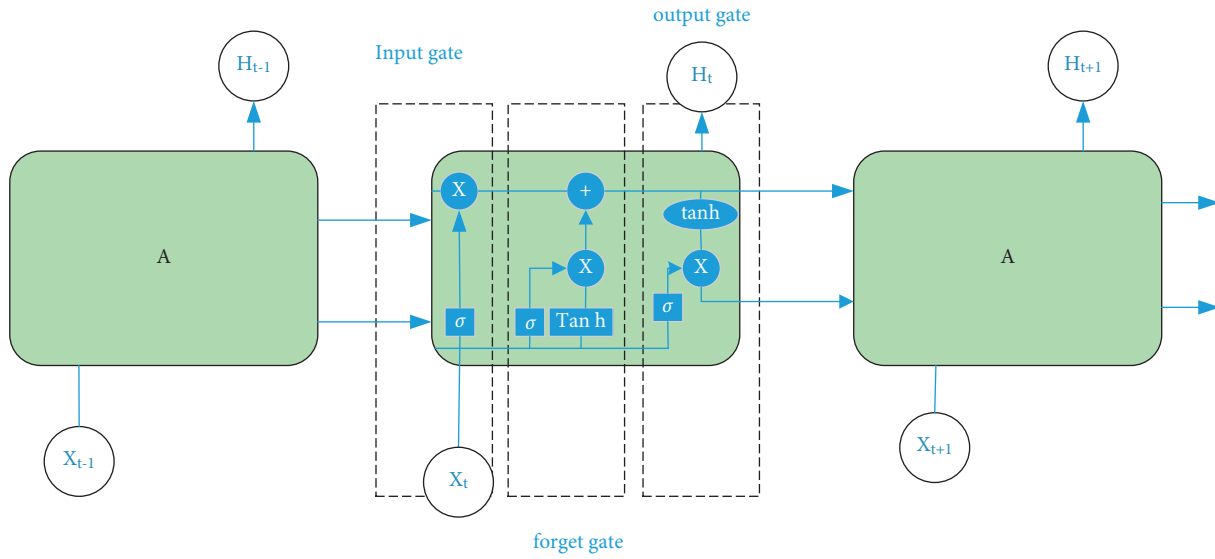


FIGURE 2: LSTM network structure diagram.

3. Recommendation Algorithm Model Construction Based on MOOC Platform Learner Model

3.1. *MOOC Platform Learner Model Construction.* Constructing the learner model mainly includes two parts: the collection of learner behavior characteristics and the construction of the learner model.

- (1) Collection of learner behavior characteristics: before establishing a learner model, learner behavior characteristics need to be extracted. The model mainly includes learner's personal information characteristics and learner's learning characteristics, as shown in Table 2.
- (2) MOOC learner model construction: the personal characteristics of learners mainly include registration information, learning ability, and historical learning content. The learning characteristics of learners include learning courses, learning time, learning experience, learning scores, learning progress, number of homework submissions, course scores, and learning ability [15–18].

Taking Figure 3 as an example, combined with the above analysis, define U as a learner, the preference obtained by analyzing the registration information is $C = \{C_1, C_2, \dots, C_n\}$, the learning course is $K = \{K_1, K_2, \dots, K_n\}$, and the learning ability is l , the learning time is h , the learning experience is e , the learning score is s , and the learning progress is p .

3.2. *Overall Description of Recommendation Algorithm Based on MOOC Platform Learner Model.* The overall structure of the learner model of the MOOC platform is shown in Figure 4. The recommendation algorithm based on the learner model mainly solves two problems. The algorithm is described as follows:

- (1) At the new user level, the cold start is alleviated by using the neural network to automatically extract user interests, and the user's favorite course category is obtained through the user's interest, and then the learning ability similarity between users is calculated within the category to obtain similar users for recommendation;
- (2) At the old user level, the user category is determined through user history information, and the learner's learning behavior characteristics and learning ability similarity are introduced for intraclass calculation to obtain the model similarity score. On this basis, the intraclass user score similarity is calculated, and finally, fusion calculation model similarity score and scoring similarity score are calculated to get the final similar users to recommend [19, 20].

3.3. *Using Neural Networks to Alleviate Cold Start.* The LSTM_Max-pooling text classification method is used to analyze the registration information of new users to obtain the interest categories of the new users and provide a basis for the next resource recommendation.

3.3.1. *Implementation of MOOC Platform Text Classification Based on LSTM_Max-Pooling.* This article uses the LSTM_Max-pooling classification model shown in Figure 5. The network model unit consists of 5 layers. The first layer is the embedding layer for word segmentation and vectorization operations; the second layer is spatial_dropout1d to improve the independence between features; the third layer is an LSTM layer, which contains 100 neurons. The input and output activation functions use \tanh , and the "gate" activation function uses sigmoid; the fourth layer uses Max-pooling to obtain the most salient features and make the model suitable for short text classification; the fifth layer is a fully connected layer using Softmax for classification. In

TABLE 2: Learner characteristics.

	Personal information	Learning characteristics
New user	Registration message	Learning ability, learning time Learning experience, learning score
Old users	History study information	Study progress number of homework submissions Course score, learning ability

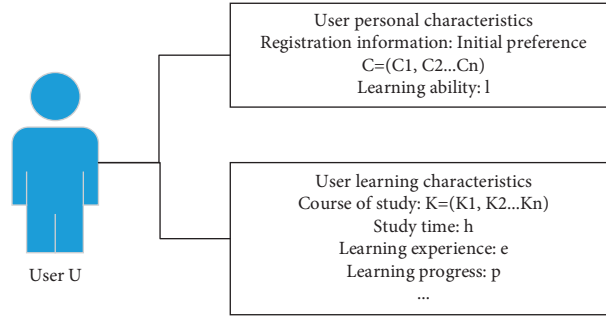


FIGURE 3: Diagram of the learner model.

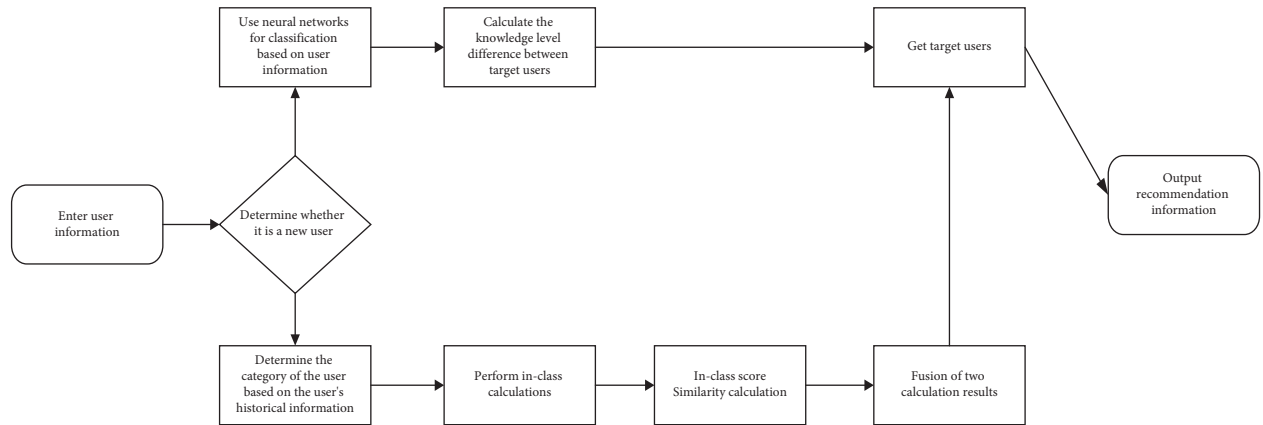


FIGURE 4: The overall flowchart of the learner model.

order to prevent the occurrence of overfitting, dropout is used in the training process and the generalization ability of the model is improved. The model workflow is as follows:

Step 1: complete sentence segmentation and vectorization

Step 2: take the word vector as input, and output the vector h_1, h_2, \dots, h_n of the hidden LSTM neural unit at n times through the LSTM layer processing

Step 3: input the vector output from the upper layer for maximum pooling operation to the feature vector h

Step 4: the upper layer feature vector h is processed by the Softmax layer to complete the text information classification [21, 22]

Definition 1. C represents the category of course resources, $C \in \{C_1, C_2, \dots, C_n\}$, which maps the user's interest category according to the course category; L represents the learning ability level of the learner; $L \in \{1, 2, \dots, 3\}$, 1, 2, 3 correspond to elementary, intermediate, and advanced levels, respectively. Newly registered users are beginner users, and registered users are evaluated according to the difficulty classification in their history learning courses; U means all users, U_{Cj} means the j -th user in category C , and U_{Ct} means target users. The calculation of the similarity within the new user class is reflected by the similarity of the learning level, and the similarity of the calculated learning level can be expressed using a standard Gaussian distribution, as shown in the following formula:

$$\text{sim}(U_{Ct}, U_{Cj})_{\text{level}} = \frac{1}{2\pi} e^{-((U_{CtL} - U_{CjL})/2)}. \quad (5)$$

3.3.2. Calculating the Similarity of the Learning Ability of Users on the MOOC Platform

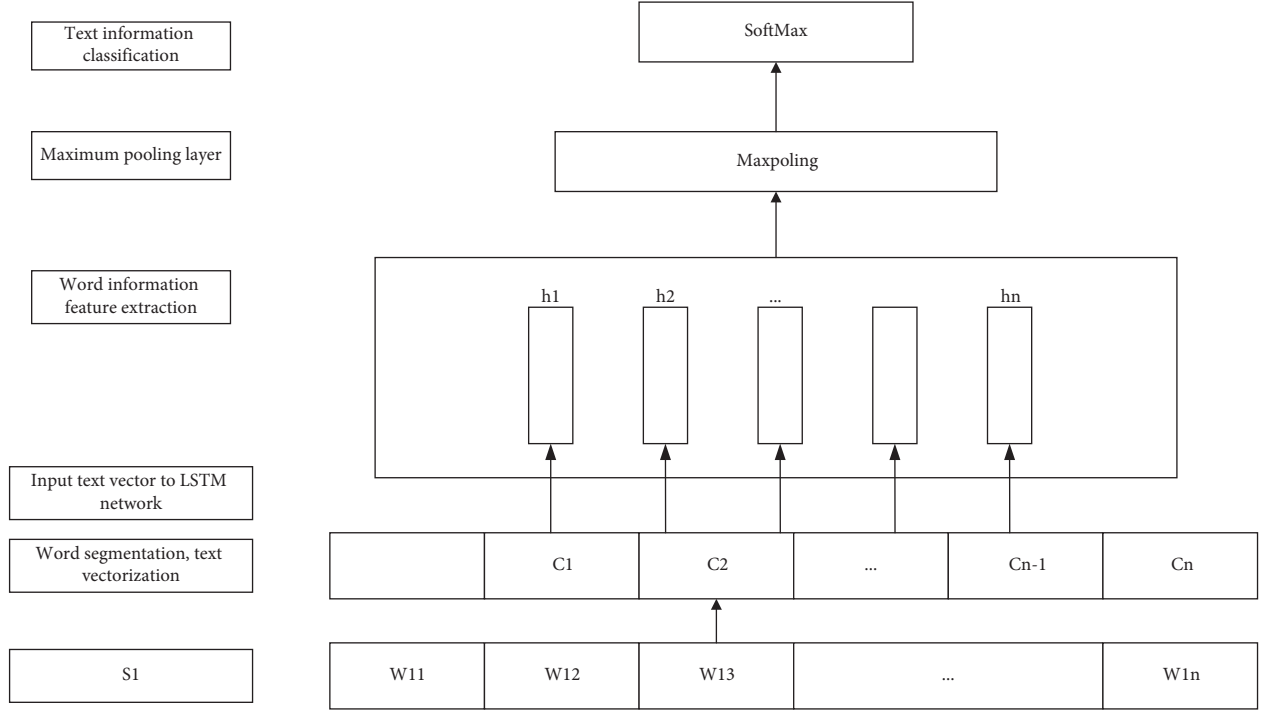


FIGURE 5: LSTM_Max-pooling classification model diagram.

3.4. Improving the Similarity Calculation Using the Learner Characteristics of the MOOC Platform

3.4.1. *Extracting Interest Tags of Old Users of the MOOC Platform.* According to the historical information of the user, the category of all the learning resources browsed by the user is obtained, and the historical learning information of the user is calculated by the statistical method to obtain the user's interest tag. The calculation method is expressed as

$$\text{hoby}(U_t) = \frac{U_{\text{th}C_i}}{U_{\text{th}}} i \in \{1, 2, \dots, n\}. \quad (6)$$

In formula (6), $U_{\text{th}C_i}$ represents the number of times the user U_t has learned the i -th type of course in the historical information, and U_t represents the total number of times the user has learned.

3.4.2. *In-Class Calculation of Learner Model on MOOC Platform.* Calculate the user interest category based on the user interest tags extracted in the previous step. The specific process is as follows:

(1) *Using the Learning Features of Old Users in Table 2 to Calculate the Similarity of Learning Features.* U represents all users, $U_{ci, t}$ represents the target user, and $U_{ci, j}$ represents other users in the same category, where $t \in U, j$ belongs to U . Use two users to calculate the behavior characteristics of the same course, $U_{ci, t} \rightarrow K$ represents the behavior characteristics of user $U_{ci, t}$ on course K , $U_{ci, j} \rightarrow K$ represents the behavior characteristics of user $U_{ci, j}$ on course K , K is for the courses that two users have studied together, and the behavior of the users on the courses is represented by a row vector, which is expressed as

$$\begin{aligned} U_{ci, t} &\rightarrow K [K_h, K_e, K_s, K_p, K_z], \\ U_{ci, j} &\rightarrow K [K_h, K_e, K_s, K_p, K_z]. \end{aligned} \quad (7)$$

In formula (9), $K_h, K_e, K_s, K_p,$ and $K_z,$ respectively, represent the proportion of the learner's time to study the course K in the total time, the proportion of the experience of obtaining the course K in the total experience, the proportion of the score of the study course K in the total score, the proportion of the progress of learning course K to the total progress, and the proportion of the number of homework submissions to the total number of times. In order to calculate the similarity of courses, use the following formula:

$$\text{sim}(U_{ci, t}, U_{ci, j})_{\text{feature}} = \frac{\sum_{i=1}^n U_{ci, j} \rightarrow K_l - U_{ci, t} \rightarrow K_l}{n}. \quad (8)$$

In formula (8), l represents the course $l \in \{1, 2, \dots, n\}$ that two users have learned together, and n represents the total number of courses that two users have studied together. Combine formulas (5) and (8) to calculate learner model similarity, which is expressed as

$$\begin{aligned} \text{sim}(U_t, U_j)_{\text{model}} &= \text{sim}(U_{ci, t}, U_{ci, j})_{\text{feature}} \\ &+ \text{sim}(U_{ct}, U_{cj})_{\text{level}} \end{aligned} \quad (9)$$

(2) *Calculation of Similarity of User Resource Scores.* Calculate the category of the target user's perceptual interest in the target user's historical information, find other users with the same interest as the target user, and use the user ID as the row index and the course ID as the column index to construct a "user-rating" table. Each row in the scoring table

represents user's rating information for all courses, each column represents the user's rating obtained by a course, the data in the table indicates the specific rating, and the "user-rating" table is shown in Table 3.

In Table 3, the user set is $U \in \{U_1, U_2, \dots, U_n\}$, and the course resource set $I \in \{I_1, I_2, \dots, I_n\}$ is used to calculate the similarity of scores between users. $\text{sim}(U_t, U_i)_{\text{score}}$ represents users U_t and U_i . The similarity of i represents the courses that are jointly graded, and $\overline{R_{U_t}}$ and $\overline{R_{U_i}}$, respectively, represent the mean value of user U_t and user U_i in the common scoring item.

$$\text{sim}(U_t, U_i)_{\text{score}} = \frac{\sum_{i \in U_t, U_j} (R_{t, i} - \overline{R_t})(R_{j, i} - \overline{R_j})}{\sqrt{\sum_{i \in U_t} (R_{t, i} - \overline{R_t})^2} \sqrt{\sum_{i \in U_j} (R_{j, i} - \overline{R_j})^2}} \quad (10)$$

3.5. Using Recommendation Algorithm Based on MOOC Platform Learner Model for Recommendation Learning. Through the above description and derivation, the learner model similarity and the score similarity are finally calculated by fusion and addition, which is specifically expressed as

$$\text{sim}(U_t, U_i)_f = \text{sim}(U_t, U_i)_{\text{model}} + \text{sim}(U_t, U_i)_{\text{score}} \quad (11)$$

According to the calculation result of formula (11), sort from high to low, and take the first N as the neighbor set. Draw the direct neighbor user set, predict the scores of the courses for which the target user does not produce scores among the final neighbor users, select the Top- N courses with the highest predicted scores for recommendation, and the calculation is expressed as

$$P_{U_t, i} = \overline{R_{U_t}} + \frac{\sum_{U_j \in s(U_t)} (R_{U_j, i} - \overline{R_{U_j}}) \cdot \text{sim}(U_t, U_j)}{\sum_{U_j \in s(U_t)} [\text{sim}(U_t, U_j)]} \quad (12)$$

In formula (12), $P_{ut, i}$ represents the predicted score of user U_t on the unrated item i , $\overline{R_{U_j}}$ represents the average score of user U_j , $R_{U_j, i}$ represents U_j 's score on item i , and $s(U_t)$ represents the neighbor set of user U_t .

4. Experiment and Result Analysis

4.1. Algorithm Evaluation Index. For details of the algorithm evaluation index, refer to the evaluation method and index of the algorithm in 2.3. This experiment uses the accuracy rate, recall rate, and F1 comprehensive evaluation index of the Top- N recommended method as the measurement standards.

4.2. Designing a Comparative Experiment. The experiment mainly includes two comparative experiments, (1) using different natural language processing methods to classify new users; (2) using traditional recommendation methods on the data set of this article to conduct comparison experiments with the recommendation algorithm proposed in this chapter.

TABLE 3: User-rating table.

	I_1	I_2	I_3	I_4	I_5
U_1	2	5	7	1	3
U_2	1	1	3	2	4
...

Experiment 1. The influence of different natural language processing algorithms on the accuracy of user interest extraction. The experiment uses the DataALL data set, which includes 50,000 users and 861 courses. The data set includes the explicit behavior information and implicit behavior of users on the courses. Information is used to verify the accuracy of the algorithm mentioned in Chapter 3 and the accuracy of the three network models of CNN, GRU, and LSTM under the same data size.

By preprocessing the data set, null values and invalid data are removed. Finally, the personal registration information of 1000 users is selected in the data set, and the experiment process is as follows:

- Step 1: import the data set, and process the data to remove stop words
- Step 2: word segmentation, vectorize the experimental data completed by word segmentation
- Step 3: vectorize the category labels
- Step 4: divide the training and test data sets
- Step 5: build an algorithm model for experiment

Experiment 2. The recommendation algorithm based on the learner model and the traditional collaborative filtering recommendation algorithm on the DataALL dataset comparative experiment.

The experiment uses the DataALL dataset to compare the accuracy, novelty, and F1 comprehensive indicators of the traditional collaborative filtering recommendation algorithm and the recommendation algorithm based on the learner model proposed in this chapter. The experimental data is randomly divided into a training set and a test set for experiments. For the convenience of calculation, Top- N is set to 10. The experimental process is as follows:

- Step 1: analyze the user's historical information to obtain the user's historical interest tag
- Step 2: according to the interest tag, find the corresponding set of the same user
- Step 3: according to the user's learning ability level similarity score and the user's learning feature similarity score, the learner model similarity score is obtained
- Step 4: get the user set from Step 1 to form the "user-course" rating table, and calculate the rating similarity
- Step 5: according to the final similarity scores between users, the final similar user set of target users is obtained
- Step 6: use the unrated items of the target user in the set of similar users of the target user to predict and output the recommended results

4.3. Analysis of Experimental Results

4.3.1. *Analysis of the Results of Experiment One.* The accuracy of different algorithms on the data set in this paper is shown in Table 4.

According to the experimental results in Table 4, as the proportion of training data increases, the accuracy of various control methods has improved. Among them, the accuracy of CNN is low compared with other methods; the overall performance of GRU method is better than CNN; and when the proportion of training set is low, the accuracy of LSTM method is greatly improved compared with the previous two methods, but with the gradual increase in the proportion of training data, the accuracy of the method in this paper has improved more obviously. It can be seen that the method in this paper has better algorithm performance for application scenarios with a large amount of data. When applied to the field of education resource management, it can alleviate the cold start problem.

4.3.2. *Analysis of the Results of Experiment Two.* Figure 6 shows the accuracy results of the comparison experiment. The higher the accuracy, the better the resources recommended to the user in line with the user's interest. The dotted line in the figure represents the accuracy of the recommendation algorithm based on the learner model at various levels, and the solid line represents the accuracy of the traditional collaborative filtering algorithm at various levels. From the experimental results, it can be seen that in the case of relatively small training data the model in this article has too many conditions to filter out, and it is difficult to find users who have too many intersections with the target user, which leads to a low recommendation accuracy. However, as the data continues to increase, the accuracy of the algorithm in this article is on an upward trend. To a certain extent, the accuracy rate is better than the traditional collaborative filtering algorithm.

Figure 7 is a graph of the novelty of the learner model recommendation algorithm. It can be seen from the graph that the method proposed in this paper is lower in novelty index than the traditional collaborative filtering algorithm in the case of a small number, but as the amount of data increases, the novelty of the recommendation method based on the learner model is gradually higher than that of the traditional collaborative filtering algorithm, which proves that the model in this paper has more advantages in recommending novelty than the traditional collaborative filtering in the big data scenario.

Since the accuracy rate only evaluates the learner model method from one aspect, in order to evaluate the comprehensive index of the algorithm based on the learner model, the $F1$ comprehensive evaluation index is used. It can be seen from Figure 8 that the $F1$ index is used in this article with less data. The proposed method is lower than the traditional collaborative filtering algorithm, but as the amount of data increases, the overall trend is increasing. The change curve of $F1$ of the recommendation method based on the learner model is gradually higher than that of the

TABLE 4: Experimental results.

Test : train	CNN	GRU	LSTM	OURS
2 : 3	0.702	0.821	0.830	0.834
3 : 7	0.710	0.824	0.832	0.840
1 : 4	0.768	0.837	0.853	0.881

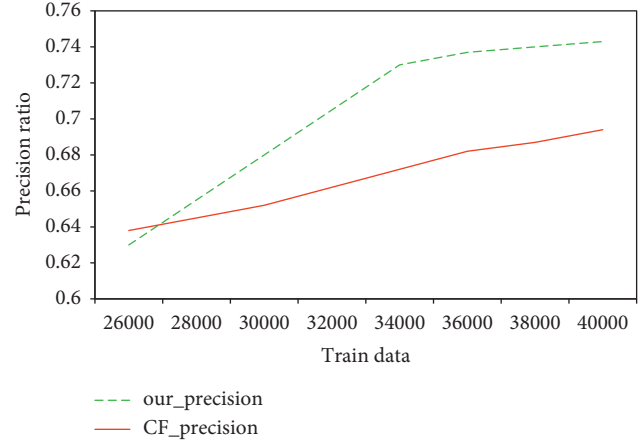


FIGURE 6: The accuracy of the learner model recommendation algorithm.

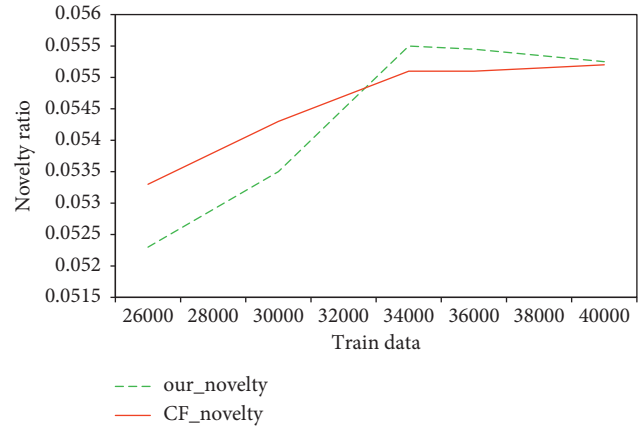


FIGURE 7: The novelty graph of the learner model recommendation algorithm.

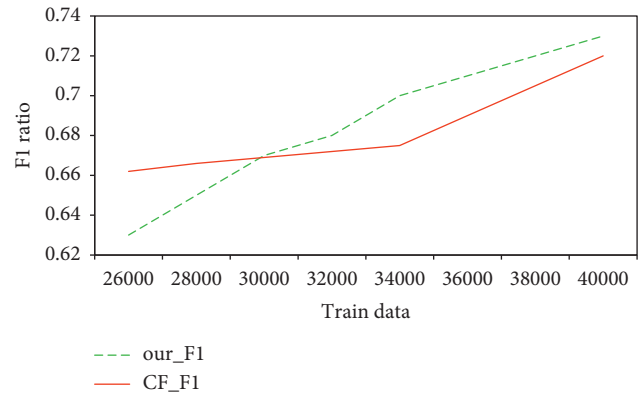


FIGURE 8: $F1$ comprehensive index diagram of learner model recommendation algorithm.

traditional collaborative filtering algorithm. It proves that the model in this paper is more suitable for scenarios with larger data.

Based on the above analysis, the recommendation method based on the learner model proposed in this paper has achieved significant results, alleviating the cold start problem of new users in the traditional recommendation system and introducing in-class calculations to improve the accuracy of resource recommendation for old users. At the same time, the model uses the objective characteristics of learners to calculate, which solves the problem of traditional collaborative filtering single score calculation.

5. Conclusion

This chapter mainly introduces the dynamic changes of learners' interest based on the MOOC platform. First, it introduces the introduction of dynamic changes in interest and the description and ideas of the recommendation algorithm; then it shows the overall structure of the algorithm and then briefly explains the structure; on this basis the implementation steps of the algorithm model are explained in detail. The algorithm mainly includes a combination of three situations. Tourist users directly perform popular resources and the latest resource fusion recommendation learning; new users use natural language processing to determine the category and then perform similarity calculations. Finally, calculate the scores of the resources that the target user may like based on the historical information of similar users and then recommend learning for them; the old users make recommendations by using the user's historical information. The difference is that this algorithm mainly focuses on the user's recent interests; thereby improving it can reflect the impact of short-term interest on the recommendation result, calculate the user's neighbor users through interest categories, learning capabilities, and resource ratings, and generate recommendation learning based on the target user's neighbor user information. Finally, the feasibility of the algorithm in this chapter is verified through design comparison experiments. It plays a more important role in the teaching management of the entire MOOC platform, but this algorithm also has certain limitations. Generally, it can only reflect short-term interests, and short-term interests will also affect the recommendation results.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] I. Barjasteh, R. Forsati, D. Ross, A.-H. Esfahanian, and H. Radha, "Cold-start recommendation with provable guarantees: a decoupled approach," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 6, pp. 1462–1474, 2016.
- [2] H. Imran, M. Belghis-Zadeh, T.-W. Chang, F. Kinshuk, and S. Graf, "PLORS: a personalized learning object recommender system," *Vietnam Journal of Computer Science*, vol. 3, no. 1, pp. 3–13, 2016.
- [3] P. Tomeo, I. Fernández-Tobías, I. Cantador, and N. D. Noia, "Addressing the cold start with positive-only feedback through semantic-based recommendations," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 25, no. Suppl. 2, pp. 57–78, 2017.
- [4] S. A. Selmane, O. Boussaid, and F. Bentayeb, "Towards collaborative multidimensional query recommendation with triadic association rules," *Decision Support Systems*, vol. 7, 2015.
- [5] J. Bhavithra and A. Saradha, "Personalized web page recommendation using case-based clustering and weighted association rule mining," *Cluster Computing*, vol. 22, no. 4, pp. 1–12, 2018.
- [6] R. Campos, R. P. D. Santos, and J. Oliveira, "Web-based recommendation system architecture for knowledge reuse in mOOCs ecosystems," in *Proceedings of the 2018 IEEE International Conference on Information Reuse and Integration*, pp. 193–200, Salt Lake, UT, USA, July 2018.
- [7] G. Manogaran, R. Varatharajan, and M. K. Priyan, "Hybrid recommendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system," *Multimedia Tools And Applications*, vol. 77, no. 4, pp. 4379–4399, 2018.
- [8] A. Gordillo, E. Barra, and J. Quemada, "A hybrid recommendation model for learning object repositories," *IEEE Latin America Transactions*, vol. 15, no. 3, pp. 462–473, 2017.
- [9] N. Akhtar and A. Mian, "Threat of adversarial attacks on deep learning in computer vision: a survey," *IEEE Access*, vol. 6, pp. 14410–14430, 2018.
- [10] E. Nachmani, E. Marciano, L. Lugosch, W. J. Gross, D. Burshtein, and Y. Be'ery, "Deep learning methods for improved decoding of linear codes," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 119–131, 2018.
- [11] S. Singaravel, J. Suykens, and P. Geyer, "Deep-learning neural-network architectures and methods: using component-based models in building-design energy prediction," *Advanced Engineering Informatics*, vol. 38, pp. 81–90, 2018.
- [12] F. Mili and M. Hamdi, "A hybrid evolutionary functional link artificial neural network for data mining and classification," in *Proceedings of the International Conference on Sciences of Electronics*, Sousse, Tunisia, March 2012.
- [13] H. T. Nguyen, C. T. Nguyen, P. T. Bao, and M. Nakagawa, "Preparation of an unconstrained vietnamese online handwriting database and recognition experiments by recurrent neural networks," in *Proceedings of the International Conference on Frontiers in Handwriting Recognition*, Shenzhen, China, October 2016.
- [14] J. Choi, S. Ryu, K. Lee et al., "ASR independent hybrid recurrent neural network based error correction for dialog system applications," in *Proceedings of the Multimodal Analyses Enabling Artificial Agents in Human-Machine Interaction*, 2014.
- [15] M. L. Alomar, E. S. Skibinsky-Gitlin, C. F. Frasser, and V. Canals, "Efficient parallel implementation of reservoir computing systems," *Neural Computing & Applications*, vol. 32, no. 3–4, pp. 1–15, 2018.
- [16] G. Çalışkan, I. Müller, M. Semtner et al., "Identification of parvalbumin interneurons as cellular substrate of fear memory persistence," *Cerebral Cortex*, vol. 26, no. 5, pp. 2325–2340, 2016.

- [17] M. K. Khribi, M. Jemni, and O. Nasraoui, "Automatic recommendations for E-learning personalization based on web usage mining techniques and information retrieval," *Journal of Educational Technology And Society*, vol. 12, no. 4, pp. 30–42, 2009.
- [18] B. Batouche, A. Brun, and B. Anne, "Unsupervised machine learning based recommendation of pedagogical resources," in *Proceedings of the Open Learning and Teaching in Educational Communities*, 2014.
- [19] J. K. Tarus, Z. Niu, and G. Mustafa, "Knowledge-based recommendation: a review of ontology- based recommender systems for E-learning," *Artificial Intelligence Review*, vol. 50, no. 1-28, 2017.
- [20] Y. Zhou, C. Huang, Q. Hu, J. Zhu, and Y. Tang, "Personalized learning full-path recommendation model based on LSTM neural networks," *Information Sciences*, vol. 444, pp. 135–152, 2018.
- [21] H. Lee, Y. Ahn, H. Lee, S. Ha, and S. G. Lee, "Quote recommendation in dialogue using deep neural network," in *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, no. 7, pp. 957–960, New York, NY, USA, July 2016.
- [22] A. Krizhevsky and I. Sutskever, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, December 2012.

Research Article

The Design of Adolescents' Physical Health Prediction System Based on Deep Reinforcement Learning

Hailiang Sun¹ and Dan Yang² 

¹School of Physical Education, Shenyang Sport University, Shenyang, Liaoning 110102, China

²Sports Department, Suqian University, Suqian 223800, Jiangsu, China

Correspondence should be addressed to Dan Yang; 16113@squ.edu.cn

Received 7 December 2021; Revised 17 December 2021; Accepted 28 December 2021; Published 29 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Hailiang Sun and Dan 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.

According to the general recognition in the first half of the last century, hypertension was not considered a kind of disease, but was regarded as a compensatory response commonly seen in the elderly, and it would not occur to younger people. Because of this erroneous cognition, many young patients fail to pay attention to their own hypertension, fail to take correct and standardized treatment, and suffer from a series of complications caused by hypertension. This article summarizes the relevant factors that affect the patient's future blood pressure from three directions: the basic characteristics of adolescent patients, the way they lower blood pressure, and the impact of the external environment. In order to make the model better fit the continuous data in the feature set of adolescents with hypertension, the structure of the internal components of the deep confidence network is optimized. Gaussian noise is introduced into the visible and hidden layers of the internal components of the network so that the stored information of the network changes from discrete to continuous during operation and improves the prediction accuracy of the blood pressure prediction model for adolescents with hypertension.

1. Introduction

The current health problems are getting younger and younger, with the emergence of adolescent hypertension becoming more serious. In the past, the demand of "getting the disease first, then treatment" has gradually changed to the demand of "early disease screening and prevention, early detection and early treatment, and personalized diagnosis"; that is, medical services are required to detect, analyze, and evaluate the health of individuals in a timely manner, provide personalized health consultation and guidance, and prevent problems before they occur. However, the current limited medical resources and the level of medical services are still unable to meet the increasing needs of people, which may lead to aggravation and deterioration of the doctor-patient relationship. With the continuous emergence of research results of deep learning technology in disease prediction, which has accelerated the application of

intelligent disease-assisted diagnosis systems in clinical practice, how to effectively ensure the health prediction of adolescents has become the focus of attention in this field. Early auxiliary diagnosis methods use traditional machine learning techniques such as Bayesian classifiers and artificial neural networks. However, traditional research methods have the limitations of insufficient feature extraction. Algorithms often extract lesion information in specific areas of the image as features, which cannot be used more comprehensively. The information predicts disease conditions. In the face of these massive physical health data, in order to obtain deeper information and dynamic monitoring of students' physical health, as well as the research and analysis of the relationship between disease and physical health, deep learning methods must be used. The core of deep learning is to extract previously unknown, valuable, and implicit information from a large amount of raw data, as well as the processing process that this technology can be understood

by people. Therefore, the level of health of adolescents in my country will be developed with new tools and new technologies, and further analysis will be conducted to discover the hidden disease information behind the physical health data, which will help early diagnosis and early warning of common hypertension in adolescents and guardianship [1–5].

2. Related Theoretical Methods

2.1. Influencing Factors in Adolescents with Hypertension. Since people have reversed their perception that hypertension is no longer a harmless compensation phenomenon but a disease, researchers have focused on the gender, age, body mass index, hypertension, and related diseases between adolescents with hypertension and healthy people. Extensive investigations and studies have been carried out on the differences in basic characteristics such as family history.

2.1.1. Gender. The data show that there are obvious differences in blood pressure of different gender groups. Although there is no research that can theoretically prove the internal connection between gender and blood pressure, the data from the five national censuses of hypertension in my country all show that gender is related to blood pressure. Professor Reckelhoff gave a more specific expression of the relationship between gender and blood pressure based on long-term monitoring data: normally, the blood pressure of men is higher than that of women of the same age. Moreover, the blood pressure of both men and women rises with age. Among them, women's blood pressure increased even more, reaching 80 years of age, and the average daily blood pressure of women and men is basically the same.

2.1.2. Age. Age is also one of the factors that affects blood pressure. This is reflected in the statistical data of my country's previous national hypertension sample surveys. There are obvious differences in the proportion of hypertensive adolescent patients of different age groups. Researcher He Sen et al. gave a more detailed explanation of the relationship between blood pressure and age: overall, with age, the systolic blood pressure and diastolic blood pressure of each age group increased significantly. Separately, the increase in systolic blood pressure becomes more pronounced with age, while the increase in diastolic blood pressure is in a U-shape during the transition from the youth to the elderly.

2.1.3. Body Mass Index. Body mass index (BMI) is calculated by dividing body weight (unit: kilogram) by the square of height (unit: meter), which is an international standard used to measure the degree of body weight. Through in-depth mining of the adolescent health status follow-up data, the study found that, with the increase of the BMI index, the incidence of hypertension in overweight and obese people increased to 1.16 to 1.28 times that of healthy people.

2.1.4. Family History of Hypertension and Related Diseases. Family history of hypertension is one of the important risk factors for the onset of hypertension. In the primary follow-up management of adolescents with hypertension in my country, family history is an important management indicator. Including not only family history of hypertension, family history of coronary heart disease, family history of stroke, and family history of diabetes and other diseases closely related to hypertension are also included in the observation object.

In summary, it can be seen that there is a significant gap in the proportion of hypertensive groups among different genders and different age groups. Obese people and people with a family history of hypertension or family history of related diseases are high-risk groups of hypertension. According to current data, cardiovascular diseases are becoming younger and younger. Therefore, gender, age, BMI index for evaluating the degree of fatness, family history of hypertension, and family history of hypertension-related diseases should be included in the blood pressure prediction model of hypertensive adolescent patients [6–10].

2.2. Deep Belief Network

2.2.1. Related Principles. As the number of layers of traditional neural network models increases, the training process becomes more and more difficult, and it is easy to fall into local extremes, which leads to the prediction effect of the deep model inferior to the shallow structure. The lack of learning algorithms suitable for deep neural networks became a major and difficult problem that hindered the development of neural networks at that time. This problem was solved in 2006. Professor Hinton of the University of Toronto proposed a deep network model that is trained by layer-by-layer initialization—deep belief net (DBN). The emergence of this deep network structure makes it feasible to train deep neural networks. Due to the excellent feature learning ability and nonlinear problem processing ability of the DBN model, it has been widely used in the field of data prediction. The network structure of the traditional DBN model is shown in Figure 1. It is composed of multiple Bernoulli-Bernoulli restricted Boltzmann machines (BB-RBMs) and a linear regression layer composition. Multiple RBM units are stacked, the output of the previous RBM unit is the input of the next unit, and the last RBM unit is connected with a linear regression layer to form the deep structure of the entire network [11].

Data prediction using the DBN model is mainly divided into two stages: unsupervised learning stage and supervised learning stage. The unsupervised learning stage refers to the unsupervised training of each RBM unit in the order from top to bottom in the network structure. When all RBM units in the DBN network are trained, the unsupervised learning process of the DBN is also completed. After the unsupervised learning stage is completed, the DBN model usually uses the same error backpropagation algorithm as the BPN model to fine-tune the network parameters and complete the supervised learning of the network. The BB-RBM unit in the

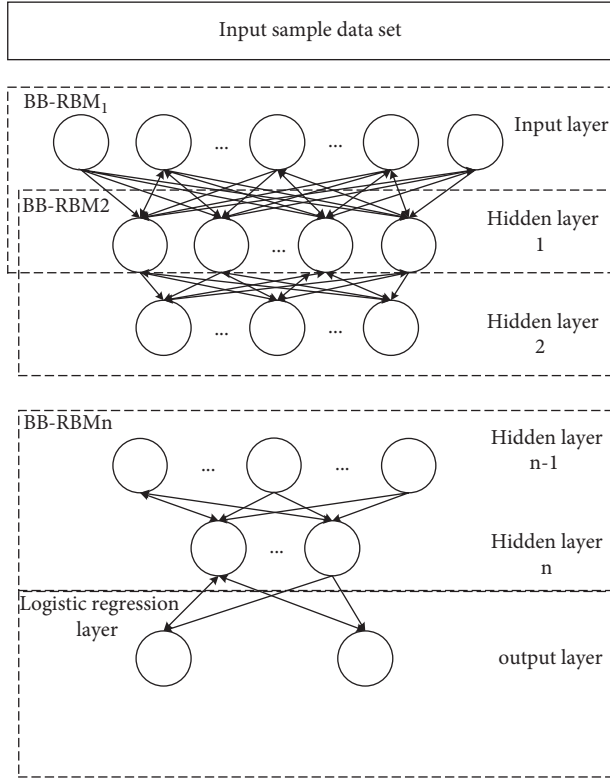


FIGURE 1: Schematic diagram of the traditional DBN model structure.

traditional DBN model is an undirected graph model with full connections between layers and no connections within the layers. The whole unit is divided into a visible layer and a hidden layer, as shown in Figure 2. The process of unsupervised training of the BB-RBM unit can be regarded as the process of extracting the characteristics of the input sample data by the DBN network. When performing unsupervised training, each BB-RBM unit is regarded as a separate system, with 1 representing the active state of its internal nodes and 0 representing the inactive state. The goal of unsupervised training is to find appropriate weights w_{ij} , visible layer bias a_i , and hidden layer bias b_j and make the input sample have the lowest energy in the system. When the system energy reaches the lowest, the current BB-RBM unit fits the data characteristics of the input sample to the greatest possible extent from the perspective of probability. Therefore, when the unsupervised training is completed, the output of the BB-RBM unit at this time can be regarded as the high-order feature representation of the input features [12–14].

2.2.2. Basic Algorithm. Taking the BB-RBM unit as shown in Figure 2 as an example, the energy it possesses can be defined as

$$E(V, H|\theta) = -\sum_{i=1}^n a_i v_i - \sum_{i=1}^n \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij} h_j. \quad (1)$$

In the formula, $\theta = \{w_{ij}, a_i, b_j\}$ represents the parameter space to be sought. According to the energy function of the

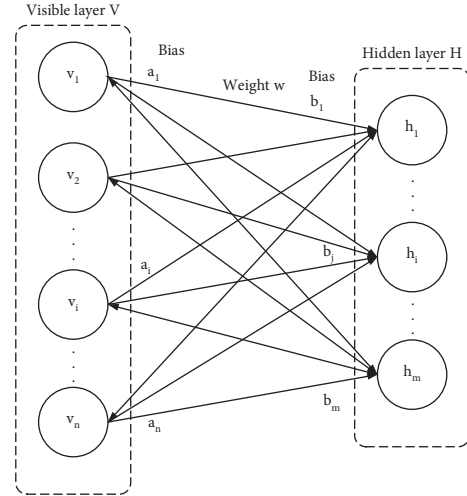


FIGURE 2: Schematic diagram of the BB-RBM structure.

BB-RBM unit, the joint probability distribution of the current state (V, H) can be determined as

$$P(V, H|\theta) = \frac{e^{-E(V, H|\theta)}}{z(\theta)}, \quad z(\theta) = \sum_V \sum_H -E(V, H|\theta). \quad (2)$$

In the formula, $Z(\theta)$ represents the partition function, which is a constant independent of the current state of the BB-RBM unit. From the joint probability distribution, the marginal probability distribution of the unit can be obtained:

$$P(V|\theta) = \sum_H P(V, H|\theta) = \frac{1}{z(\theta)} \sum_H e^{-E(V, H|\theta)}. \quad (3)$$

Equation (3) expresses the probability of the input sample appearing in the parameter space θ , and the purpose of BB-RBM unsupervised training is to find a suitable parameter space θ to maximize the probability of the input sample; therefore, its distribution is determined. This completes the training process of the current BB-RBM unit. In order to determine the distribution, the partition function $z(\theta)$ needs to be calculated according to equations (2) and (3), which requires a total of 2^{m+n} operations. In order to reduce computational complexity, Professor Hinton, the proponent of the RBM, proposed a contrastive divergence (CD) algorithm for fast training of RBM units. The specific process is as follows:

- (1) Calculate the state distribution probability of nodes in the hidden layer H_0 according to the original input V_0 of the visible layer of the current BB-RBM unit [15–19]:

$$P(h_j = 1|\theta) = \sigma\left(\sum_{i=1}^n v_i w_{ij} + b_j\right). \quad (4)$$

In the formula, $\sigma(x)$ represents the activation function. According to the state distribution probability of H_0 , perform Gibbs sampling to obtain the state information of the nodes in H_0 .

- (2) According to the state information of the nodes in H_0 , reconstruct the visible layer of the BB-RBM unit to obtain the distribution probability of the node state of the new visible layer V_1 :

$$P(v_j = 1|\theta, \theta) = \sigma\left(\sum_{i=1}^N v_i w_{ij} + b_j\right). \quad (5)$$

Similarly, the Gibbs sampling method is used to obtain the status information of the nodes in V_1 .

- (3) According to the state information of the nodes in V_1 , reconstruct the hidden layer of the BB-RBM unit to obtain the distribution probability of the node state of the new hidden layer H_1 :

$$P(h_j = 1|v, \theta) = \sigma\left(\sum_{i=1}^N v_i w_{ij} + b_j\right). \quad (6)$$

The Gibbs sampling method is used to obtain the status information of the nodes in H_1 .

- (4) Update the connection weight of the BB-RBM unit w_{ij} , hidden layer offset b_j , and visible layer offset a_i :

$$\begin{cases} w_{ij}^{t+1} = w_{ij}^t + \varepsilon(H_0 V_0^T - H_1 V_1^T), \\ a_i^{t+1} = a_i^t + \varepsilon(V_0 - V_1), \\ b_j^{t+1} = b_j^t + \varepsilon(H_0 - H_1). \end{cases} \quad (7)$$

In the formula, t represents the number of iteration operations. Iteratively update the parameter space until all samples are trained, and the unsupervised training process of the BB-RBM unit is completed.

On the whole, the DBN model is an interdisciplinary product. Its basic unit is a restricted Boltzmann machine rooted in statistical mechanics. The model itself is a neural network in the field of artificial intelligence. The advantage of the DBN model is that the energy model is introduced to simulate the data distribution so that the sample data with complex internal relationships have a general likelihood function that describes their probability distribution, making it feasible to learn the probability distribution of complex sample data. The traditional DBN model also has its limitations. First, because the BB-RBM has only two states, active and inactive, the stored information of the network neurons can only be 0 or 1, and some sample data may be lost. Second, like the BPNN model, the network structure is uncertain. The values of the number of hidden layers and the number of hidden layer nodes lack theoretical support [20].

2.3. Evaluation Index for the Accuracy of Adolescent Blood Pressure Prediction. At present, the American Association for the Advancement of Medical Devices (AAMI) standard is used internationally to evaluate whether a blood pressure meter is qualified. According to the AAMI standard, the maximum error allowed by the instrument used for blood pressure measurement is 5 mmHg. Therefore, this article also uses ± 5 mmHg as the criterion to evaluate the

prediction accuracy of the blood pressure prediction model for adolescents with hypertension; that is, when the model predicts blood pressure and the patient's actual blood pressure value is within 5 mmHg, it is judged here that the prediction result is accurate; otherwise, it is judged as inaccurate. Define it as the model prediction accuracy Acc:

$$\text{Acc} = \frac{m}{n} \times 100\%. \quad (8)$$

In the formula, n represents the sample size, and m represents the number of samples whose prediction error is within the criterion. The higher the prediction accuracy, the higher the accuracy of the model for predicting the patient's blood pressure [21, 22].

3. Establishment of a Predictive Model for Adolescents' Physical Health

3.1. Data Preprocessing. Data preprocessing is the process of processing the collected raw data into a high-quality dataset that meets the needs of research goals. Good data preprocessing can effectively improve the efficiency of subsequent data modeling and data analysis. This article analyzes the influencing factors of the blood pressure of adolescent patients, conducts a series of data preprocessing works on the collected outpatient follow-up original records and weather original records of adolescent patients with hypertension, and establishes a high-quality dataset of characteristics of adolescent patients with hypertension. The overall process of the work is shown in Figure 3 [23].

3.2. Experimental Environment. The experiment uses a personal computer, and the detailed configuration is shown in Table 1.

3.3. Adolescents' Physical Health Prediction Model Based on the DBN

3.3.1. Predictive Model Process. For the DBN-based blood pressure prediction model for adolescents with hypertension, its hidden layer structure is also related to the final prediction accuracy of the model. Moreover, because the DBN model has one more unsupervised learning process than the BPNN model, the selection process of the network structure of the DBN model is often more complicated than that of the BPNN model. At present, there is still no complete theoretical basis for the selection of the number of hidden layers of the DBN model and the number of nodes in each hidden layer. Therefore, this article sets up multiple groups of controlled trials to determine the optimal number of nodes from top to bottom and establish a blood pressure prediction model based on the DBN for adolescents with hypertension.

3.3.2. Model Performance Evaluation. The overall model construction process is shown in Figure 4. The specific steps are as follows:

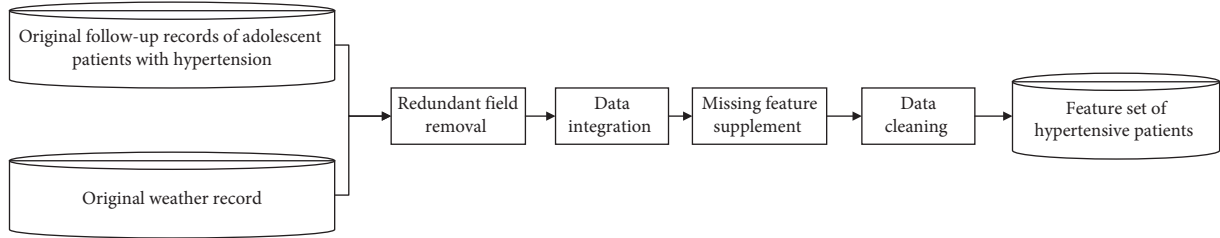


FIGURE 3: Data preprocessing process.

TABLE 1: Experimental environment configuration.

Software and hardware	Configuration
CPU	Intel Core i7
RAM	32 G RAM
Hard disk	1 TB HDD
Operating system	Windows 10
Programming language and version	Python 3.64
Development environment	PyCharm integrated development environment

- (1) Randomly select 10,000 records in the characteristic dataset of adolescent patients with hypertension as the training sample set of the DBN model, and the remaining 3762 records are used as the test sample set of the DBN model.
- (2) Considering the performance of the experimental equipment and the risk of overfitting, the maximum number of hidden layers of the DBN hypertensive adolescent blood pressure prediction model is set to 3; the value set of the number of hidden layer nodes \emptyset is set to $\{10, 20, 30, 40, 50, 60\}$; the initial value is set to 10; the maximum number of iterations for supervised training is set to 1000; the maximum number of iterations for unsupervised training is set to 5000.
- (3) Initialize the RBM unit. Set the learning rate of its unsupervised training $e1=0.5$; the activation function uses the sigmoid function; the visible layer bias a_i , the hidden layer bias b_j , and the connection weight of the hidden layer and the visible layer w_{ij} are initialized to the interval $(0, 0.1)$ value in range.
- (4) Randomly extract the training sample set samples, and unsupervised training of the current RBM unit is carried out according to the CD algorithm.
- (5) After reaching the maximum number of iterations, the unsupervised training ends. Construct a logistic regression layer with 2 nodes as the output layer of the network, which are used to predict diastolic blood pressure and systolic blood pressure, respectively. Connect the regression layer node to the hidden layer node of the current RBM unit, and initialize the connection weight between the hidden layer and the output layer w_{ij} and the bias b_j corresponding to each node in the output layer to be in the interval $(0, 0.1)$. For random values, perform the supervised training of the DBN model according to the same error backpropagation algorithm as the BPNN model, and update the network parameters w_{ij} , b_j , w'_{ij} , and b'_j .
- (6) After the supervised training is completed, input the test sample dataset into the DBN-based blood pressure prediction model for adolescent hypertensive patients, and record the blood pressure prediction results of the DBN model under the current structure for the hypertensive adolescent patients in the test set.
- (7) Take the set value of the number of nodes in the next hidden layer in the set \emptyset , and repeat Steps 3 to 6, until all the values in \emptyset are traversed. Establish a DBN hypertensive adolescent blood pressure prediction model with the same number of hidden layers and a different bottom RBM unit structure, and determine the bottom RBM unit structure with the best prediction effect of the DBN model under the fixed number of hidden layers according to the prediction results, which is the current optimal structure of the RBM unit.
- (8) Add a hidden layer, and judge whether the number of added hidden layers is greater than the preset maximum value. If it is not exceeded, remove the logistic regression layer of the DBN model, add a new RBM unit, and use the hidden layer of the previous unit as the visible layer of the new unit. The initial value of the number of nodes in the hidden layer of the new unit is also set to 10. Return to Step 3, and establish DBN models with different hidden layer structures; if it exceeds, compare the results of the DBN model predicting the patient's blood pressure under different hidden layer structures, and select the model with the highest prediction accuracy. Finally, according to the above steps, the prediction

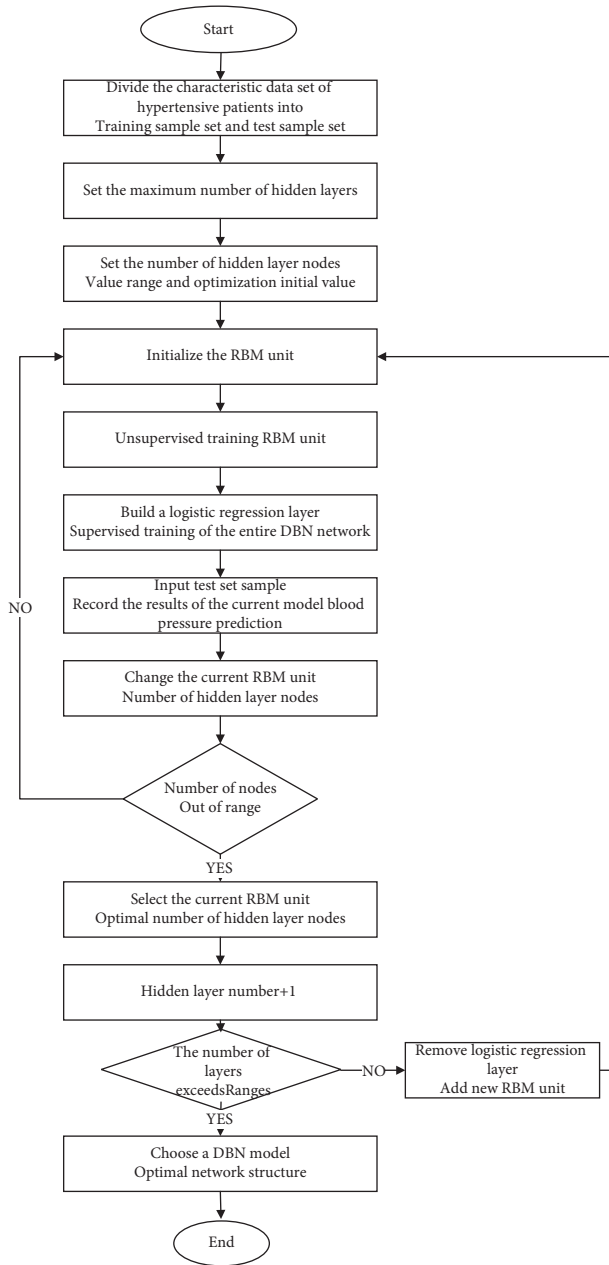


FIGURE 4: The process of constructing a predictive model for adolescents with hypertension based on the DBN.

results of the DBN model for the patient's diastolic and systolic blood pressure under different hidden layer structures are obtained, as shown in Table 2.

From the experimental results in Table 2, it can be seen that when the hidden layer of the DBN-based hypertensive adolescent blood pressure prediction model is changed from one layer to two layers, the prediction accuracy of systolic blood pressure and diastolic blood pressure is significantly improved, indicating that, in hypertension, in the problem of blood pressure prediction in adolescent patients, the deep network has a better complex relationship mapping than the shallow network.

When the network continues to deepen and the hidden layer increases to 3 layers, the blood pressure prediction accuracy rate of the DBN model decreases. This may be due to the accumulation of overall network errors and overfitting; when the number of layers is 2 and the number of nodes is 30 and 10, the model predicts the patient's diastolic blood pressure and systolic blood pressure with the highest accuracy, which are 85.41% and 75.84%, respectively. Therefore, the network structure of the DBN-based blood pressure prediction model for adolescents with hypertension is finally determined to be four layers, and the number of nodes in the two hidden layers is 30 and 10 from the input layer to the output layer.

4. Adolescents' Physical Health Prediction Model Based on the Improved Deep Belief Network

The fully continuous deep belief network model that uses the adaptive distance estimation algorithm for supervised learning is defined as FCDBN-Adam; the fully continuous deep belief network model that uses the stochastic gradient descent algorithm for supervised learning is defined as FCDDN-SGD; a deep artificial neural network model that has the same network structure as the above two models and uses an adaptive distance estimation algorithm for supervised learning is defined as ANN-Adam; it uses a stochastic gradient descent algorithm for supervised learning depth. The trust network model is defined as DBN-Adam. The settings of each model are shown in Table 3. Explore whether the FCDBN-Adam hypertensive adolescent blood pressure prediction model proposed in this article has higher accuracy and stability than other deep models in the face of different types of hypertensive adolescent patients.

4.1. Different Age Groups. Age is one of the important factors that affects the blood pressure level of patients. With the increase of age, the body's immunity decreases, and the body's ability to control blood pressure becomes worse. As a result, older adolescents with hypertension bear a higher risk of disease than younger patients. Considering that the blood pressure levels of hypertensive adolescents of different ages have different influences on the factors affecting their blood pressure levels, the characteristic dataset of hypertensive adolescents is now grouped by age (see Table 4 for grouping results) to explore in this article whether the established blood pressure prediction model for adolescents with hypertension can reveal the changes of blood pressure caused by the affected factors in patients of different ages.

100 samples were randomly selected from each group to form test samples, and the remaining samples were used as training samples to test the performance of each deep neural network model in Table 3 in the face of patients of different ages. The results are shown in Figures 5 and 6.

It can be seen from the prediction results of the blood pressure level of patients of different age groups in Figures 5 and 6 that the four models predict the diastolic blood

TABLE 2: The prediction results of the DBN model under different hidden layer structures.

Group no.	Hidden layers	Hidden layer structure	Systolic pressure Acc (%)	Diastolic pressure Acc (%)
1	1	10	59.9	63.9
2	1	20	61.8	70.9
3	1	30	69.1	75.8
4	1	40	68.4	74.4
5	1	50	60.1	65.9
6	1	60	53.2	66.3
7	2	30-10	75.9	85.1
8	2	30-20	71.8	85.0
9	2	30-30	70.8	78.1
10	2	30-40	70.2	78.3
11	2	30-50	69.3	76.1
12	2	30-60	64.9	73.4
13	3	30-10-10	74.9	84.7
14	3	30-10-20	73.1	83.1
15	3	30-10-30	72.9	84.2
16	3	30-10-40	67.9	80.5
17	3	30-10-50	67.9	78.0
18	3	30-10-60	67.9	75.1

TABLE 3: Settings of the experimental model.

Model name	Network structure	Internal unit	Supervised learning algorithm
FCDBN-Adam	32-40-20-10-2	GG-RBM	Adam
FCDBN-SGD	32-40-20-10-2	GG-RBM	SGD
ANN-Adam	32-40-20-10-2	Normal neuron	Adam
DBN-Adam	32-30-10-2	BB-RBM	Adam

TABLE 4: Sample distribution of the dataset of characteristics of adolescent patients with hypertension (grouped by age).

Group	Age range	Sample size
1	14-17 years old	1800
2	18-25 years old	4900

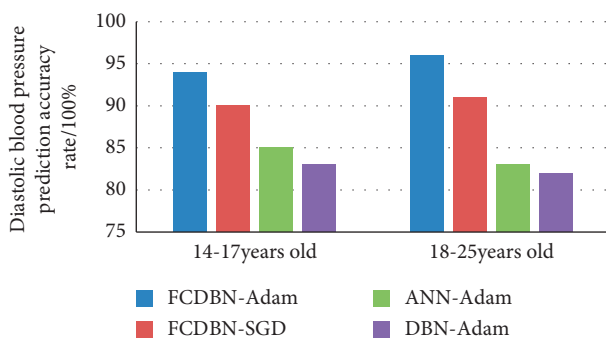


FIGURE 5: Diastolic blood pressure prediction results of different depth models (grouped by age).

pressure level of patients of different age groups under the premise that the number of training samples given by each age group is different. The above accuracy rate did not show a significant difference, but in predicting the systolic blood pressure level of patients, it may be due to the influence of a special undiscovered systolic blood pressure influencing factor. In the prediction of the pressure level, the prediction accuracy of each deep neural network model has decreased

slightly, and the FCDDN-Adam model proposed in this paper can still maintain a prediction accuracy of more than 80%, which exceeds the other three deep models.

4.2. Different Follow-Up Cycles. The follow-up physicians set up different follow-up cycles for the patients according to the severity of the condition of the adolescents with hypertension. Patients with mild illness have a longer follow-up period, while patients with severe illness have a relatively shorter follow-up period. Moreover, the follow-up period can reflect the length of time the patient continues the current antihypertensive treatment. Long-term drug interventions and lifestyle interventions are likely to change the law of influencing factors on the blood pressure level of patients. Therefore, the characteristic dataset of adolescents with hypertension is now grouped according to the length of the follow-up period. See Table 5 to explore in this article whether the established blood pressure prediction model for adolescents with hypertension can reveal the changes of blood pressure caused by the affected factors in patients with different follow-up periods.

100 samples were randomly selected from each group to form test samples, and the remaining samples were used as training samples to test the performance of the deep neural network models in Table 3 in the face of patients with different follow-up periods. The results are shown in Figures 7 and 8.

It can be seen from the prediction results of the blood pressure level of patients in different follow-up periods in

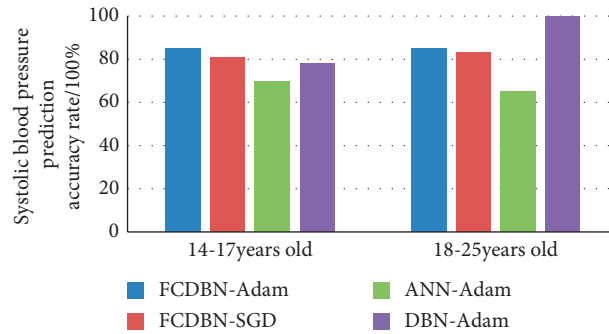


FIGURE 6: The prediction results of systolic blood pressure of different depth models (grouped by age).

TABLE 5: Sample distribution of the characteristic dataset of adolescent patients with hypertension (grouped by the follow-up period).

Group	Follow-up cycle	Number of samples
1	Within two weeks	2900
2	Two weeks to one month	5700
3	More than one month	5000

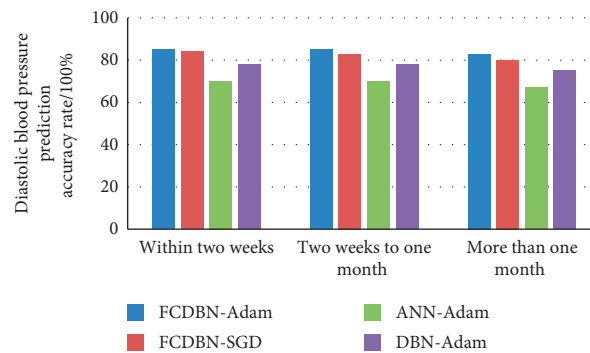


FIGURE 7: Prediction results of systolic blood pressure of different depth models (grouped by the follow-up period).

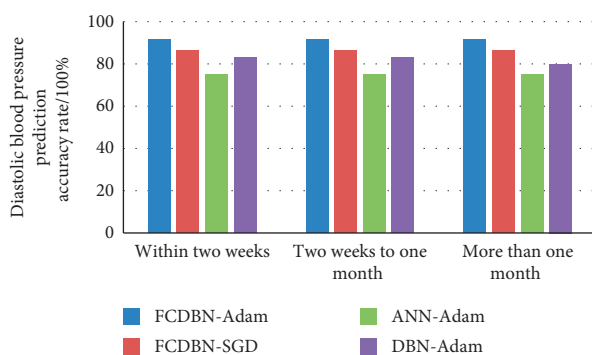


FIGURE 8: Diastolic blood pressure prediction results of different depth models (grouped by the follow-up period).

Figures 7 and 8 that, as the prediction time span increases, the prediction accuracy of each deep neural network model decreases, and the proposed method based on FCDBN-Adam is high. Compared with other models, the blood pressure prediction model for adolescent patients with blood pressure is more stable. The prediction accuracy of diastolic blood pressure is always above 90%,

and the prediction accuracy of systolic blood pressure is maintained at about 85%. It has higher prediction accuracy and can predict well.

4.3. Result Analysis. Judging from the overall results of the four models for predicting the patient's blood pressure, it can be seen that, under the same number of supervised training iterations, in terms of prediction accuracy, the FCDBN-Adam model is better than the FCDBN-SGD model. Among them, the Adam algorithm has a higher learning efficiency than the SGD algorithm; the prediction accuracy of the FCDBN-Adam model far exceeds that of the ANN-Adam model with the same structure, indicating that the deep belief network model can be better than the multilayer neural network model. Characterize the complex nonlinear relationship between the future blood pressure situation of the human body and its various influencing factors; the prediction accuracy of the FCDBN-Adam model is significantly higher than that of the DBN-Adam model, indicating that the traditional RBM unit inside the deep confidence network is replaced with a double-layer Gaussian structure. The GG-RBM unit can effectively improve the

prediction accuracy of the blood pressure prediction model for adolescents with hypertension.

To sum up, this article proposes a blood pressure prediction model based on FCDBN-Adam hypertensive adolescent patients. Although there are differences in the blood pressure prediction results of different types of hypertensive adolescent patients, the overall prediction accuracy of the model remains at a high level, and the prediction accuracy of diastolic blood pressure has reached about 85%, and the accuracy of predicting systolic blood pressure is over 90%, which is better than the other three deep neural network models. It can better characterize the complex nonlinear relationship between the future blood pressure of the human body and its various influencing factors. It is suitable as a model for predicting the blood pressure level of adolescents with hypertension under the complex influence of multiple factors.

5. Conclusion

This article first optimizes the internal units of the blood pressure prediction model for adolescents with hypertension based on the deep confidence network and uses a restricted Boltzmann machine (GG-RBM) unit with a double-Gaussian structure to replace the double-layer Bernoulli in the original model. The structure of the restricted Boltzmann machine (BB-RBM) unit has established a blood pressure prediction model for adolescents with hypertension based on the fully continuous deep confidence network (FCDBN). Experiments show that the improved model can better fit the characteristics of adolescent patients with hypertension and improve the accuracy of the model for predicting the blood pressure level of patients.

Secondly, considering the large number of input factors in the current model and the complicated selection process of the hidden layer structure, this paper designs a method based on the traditional network structure determination method to determine the number of hidden layer nodes in a fully continuous deep confidence network model based on reconstruction errors. The method of hiding the number of layers omits the step of supervised learning in the network structure selection process, saves the time cost of blood pressure prediction modeling for adolescents with hypertension, and enhances the scalability of the model.

Finally, this paper constructs four different types of deep neural network models, FCDBN-Adam, FCDBN-SGD, DBN-Adam, and ANN-Adam, respectively, and evaluates the blood pressure prediction accuracy of the four models in the face of different types of adolescents with hypertension. The experimental results show that the blood pressure prediction model for adolescents with hypertension based on FCDBN-Adam proposed in this paper performs best. In the face of different types of adolescents with hypertension, it can maintain a high prediction accuracy of patient blood pressure. The average prediction accuracy of diastolic blood pressure exceeds 90%, and the systolic blood pressure reaches about 85%. It is based on four types of deep learning. In the future, with the widespread application of the adolescent prediction system, it is believed that other adolescent diseases can also be well predicted.

Data Availability

The dataset can be accessed from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] J. Gssowski and K. Piotrowicz, "Blood pressure target: high time that we finally agreed what is healthy," *Hypertension*, vol. 68, no. 5, pp. 1103–1105, 2016.
- [2] G. A. Mensah, Z. S. Galis, L. J. Fine, D. F. Levy, and G. H. Gibbons, "Building on a legacy of hypertension research: charting our future together," *Hypertension*, vol. 69, no. 1, pp. 5–10, 2017.
- [3] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Image Net classification with deep convolutional neural networks," in *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 1097–1105, Lombok, Indonesia, September 2012.
- [5] P. Ocampo, A. Moreira, N. Coudray et al., "P1.09-32 classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning," *Journal of Thoracic Oncology*, vol. 13, no. 10, p. S562, 2018.
- [6] A. Teramoto, T. Tsukamoto, Y. Kiriya, and H. Fujita, "Automated classification of lung cancer types from cytological images using deep convolutional neural networks[J/OL]," *BioMed Research International*, vol. 2017, Article ID 4067832, 2017.
- [7] H. Chan Stephanie, C. Van Hee Victor, B. Silas, and T. Sakellaropoulos, "Long-term air pollution exposure and blood pressure in the sister study," *Environmental Health Perspectives*, vol. 123, no. 10, pp. 951–958, 2019.
- [8] L. Sarah, L. Li, S. Paul, and W. Siekmeyer, "Seasonal variation in blood pressure and its relationship with outdoor temperature in 10 diverse regions of China: the China Kadoorie Biobank." *Journal of Hypertension*, vol. 30, no. 7, pp. 1383–1391, 2012.
- [9] A. Miersch, M. Vogel, R. Gausche et al., "Influence of seasonal variation on blood pressure measurements in children, adolescents and young adults," *Pediatric Nephrology*, vol. 28, no. 12, pp. 2343–2349, 2013.
- [10] J. F. Dartigues, J. M. Lacombe, P. Ducimetiere, B. Malin, and N. Freitase, "Relationship between blood pressure and outdoor temperature in a large sample of elderly individuals: the three-city study," *Archives of Internal Medicine*, vol. 169, no. 1, pp. 75–80, 2020.
- [11] P. W. Lof, X. T. Lic, B. Malin, N. Freitase, and J. K. Wang, "Continuous systolic and diastolic blood pressure estimation utilizing long short-term memory network," in *Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1853–1856, IEEE, Piscataway, NJ, June 2017.
- [12] G. Paolo, R. Melvyn, and R. Das, "Particulate matter air pollution and ambient temperature: opposing effects on blood pressure in high-risk cardiac patients," *Journal of hypertension*, vol. 33, no. 10, pp. 2032–2038, 2015.

- [13] S. M. Davide and M. Fernando, "Nonstationary Z-Score measures," *European Journal of Operational Research*, vol. 260, no. 1, pp. 0377–2217, 2016.
- [14] R. Ge, F. Huang, and M. Fernando, "Escaping from saddle points: online stochastic gradient for tensor decomposition," *Journal of Machine Learning Research*, pp. 1–46, 2015.
- [15] D. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *Proceedings of the Published as a conference paper at the 3rd International Conference for Learning Representations*, pp. 1–15, ICLR, San Diego, May 2015.
- [16] Guido and N. Ay, "Refinements of universal approximation results for deep belief networks and restricted Boltzmann machines," *Neural Computation*, vol. 23, no. 5, 2021.
- [17] K. Swersky, B. Chen, B. Malin, and N. Freitas, "A tutorial on stochastic approximation algorithms for training restricted Boltzmann machines and deep belief Nets," *International Techniques and Application*, , pp. 80–89, IEEE, 2010.
- [18] G. E. Hinton, *A Practical Guide to Training Restricted Boltzmann Machines[R]*, Technical Report UTML TR2010-003, University of Toronto, Toronto, Canada, 2010.
- [19] A. Fischer and C. Igel, "Bounding the bias of contrastive divergence learning," *Neural Computation*, vol. 23, no. 3, pp. 664–673, 2010.
- [20] H. Lee and R. Ranganath, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representation," *Proceedings of the 26th International Conference on Machine Learning*, pp. 609–606, New York, USA, 2009.
- [21] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine learning*, vol. 2, no. 1, pp. 1–127, 2009.
- [22] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural Computation*, vol. 14, no. 8, pp. 1771–1800, 2002.
- [23] M. Zorzi, A. Testolin, and I. P. Stoianov, "Modeling language and cognition with deep unsupervised learningA tutorial overview," *Frontiers in Psychology*, vol. 4, no. 515, pp. 1–14, 2020.

Research Article

Collaborative Research on Mouth Shape and Lyrics in Singing Practice Based on Image Processing

Lujia Xu¹ and Chen Chen² 

¹Xiamen University Tan Kah Kee College, Zhangzhou 363123, China

²The First Affiliated Hospital of Xiamen University, Xiamen 361000, China

Correspondence should be addressed to Chen Chen; g20041117@xmu.edu.cn

Received 22 November 2021; Revised 10 December 2021; Accepted 14 December 2021; Published 27 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Lujia Xu and Chen 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.

Image processing is a mainstream processing method. When people enjoy artists' singing videos, there will be a problem that the subtitles of the lyrics are out of sync with the singer's mouth shape. This problem needs to be solved using image processing technology, letting the computer realize lip-reading recognition function and correct the mouth shape and lyrics subtitles in the image according to the extracted lip-reading data, so that the mouth shape and lyrics in singing practice can be synchronized. Lip-reading information can effectively improve the accuracy of language cognition, save part of capital and manpower investment, and make viewers get a good audio-visual interactive experience. The results show the following: (1) After the UI test, the system user interface function design is reasonable and there is no bad BUG. We can find that the average processing time of each frame is 628 ms, the system performance evaluation is good, and the success rate can be as high as 98.80%. 0.36724 s is the average time for each step when the system processes the image. (2) The human image can basically identify the portrait area and lip area from various angles. (3) Compared with DCT and DWT, the recognition rate of the two cascade lip region feature extraction methods is improved by nearly 10%, and the feature vector dimension is reduced by nearly 65%. (4) Classify the mouth shape more finely and optimize the image of the tester's mouth shape to make the mouth shape closer to the standard mouth shape. (5) After systematic correction of mouth shape and subtitles, the success rate is higher than 90%. Finally, we can find that the running effect is good and the method has achieved high results, which can carry out the details of the next optimization work.

1. Introduction

With the spring breeze of science and technology sweeping the world, times have quietly changed dramatically. How to use the power of science and technology to make people's daily life more efficient and convenient is the most frequently considered problem nowadays. In recent years, people are using lip-reading more and more frequently, although they do not realize that they are using it. However, lip reading technology is no longer a "spring snow." It has gone out of the laboratory and gradually stabilized from little known to development and research. The relevant application technology is becoming more and more practical, and more and more people from outside are beginning to understand and learn its technology and characteristics. In this

study, if we want to coordinate the mouth shape and lyrics in the video, the first step is to accurately locate people's lips in the image, and then a series of mouth shape recognition tasks are carried out. There are many literature and materials related to computer lip-reading recognition in existing journals, papers, and other places, and we select some of them for reference study here. Literature [1] introduces the reality and development level of lip reading, which arouses people's attention and interest. Literature [2] proposes algorithms for lip detection, feature extraction, and other technologies to realize a lip-reading prototype system. In [3], the face recognition technology is used in the attendance system. Literature [4] uses optical flow to analyze lip images, calculates two visual feature sets in each frame, and proposes a multimode speech recognition method. Literature [5]

analyzes lip-reading technology according to typical deep learning and lists the existing lip-reading databases. Literature [6] uses a three-dimensional motion capture system to extract accurate parameters of facial motion features through lip reading. Literature [7] describes the public database of lip reading, target speech content, equipment, camera orientation, frame rate, and application. In [8], the coupled hidden Markov model is used to combine audio-visual signals with Polish speech recognition under the condition of highly interfered audio signals. Literature [9] studies the dual-mode fusion algorithm of video and audio, combining lip-reading speech recognition and information fusion technology. Literature [10] tests the fusion technology of different modes and analyzes different methods of audio-visual speech recognition. Literature [11] proposed the application of a lip-reading method based on a convolution neural network to tandem three-sequence key frame images. Literature [12] developed the deep convolution neural network model of HLR-Net for lip reading. Literature [13] proposes a new VSR method, multiangle lip reading, for audio-visual speech recognition. Literature [14] classifies those binoculars, enzymes, and muscle contractions in language when lip reading using multichannel SEMG signals. Literature [15] designs a graphic structure and lip partial cut network, using a local adjacent function extractor and lip reading with multilevel function fusion.

2. Theoretical Basis

2.1. Video and Audio Bimodal Corpus. Video and audio bimodal corpus [16] serves as the research basis. It will process the video according to the speech signal and visually include various face location images and mouth images related to the speech pronunciation. Because corpora involve many factors that need to be considered, such as format, coding, environment, and storage, and because there are few bimodal corpora that can be shared in the market, there are Tulips, M2VTS, ViaVoice, and other corpus databases in the world. There are the relevant parameters of the corpus, as shown in Table 1.

2.2. Lip Detection and Positioning Method. Different lip feature extraction methods obtain different lip areas [17]. If the first step of this positioning is not successful, all the functions based on lips will not be realized smoothly. The detection of lip area requires the system to quickly and accurately find the approximate range of the face in each frame of an image. After face recognition, it is necessary to accurately locate the lips according to their characteristics. In this way, we can ensure that after the previous work, we have laid a good foundation for lip feature extraction and make the subsequent image processing smoother and easier. The following is an introduction to typical lip detection methods.

2.2.1. Face Detection. Observing the facial features of human beings, there are three main human organs: eyes, nose, and mouth. Unless there are man-made or unexpected factors,

their positions will not change much for a long time, and they have relative stability. Based on the stable features of face organs, we can use computer programs to detect the specific location of the face, and then use the Canny operator and projection method to quickly locate the lip area. W_f denotes the face width and H_f denotes the face height and the lip area of interest is as follows:

$$\begin{aligned} \frac{1}{4}W_f \leq x \leq \frac{3}{4}W_f, \\ \frac{2}{3}H_f \leq y \leq \frac{1}{15}H_f. \end{aligned} \quad (1)$$

However, due to the limitations of technology, time, and space, this detection method cannot accurately obtain and recognize the position of the face and specific lip area in some cases. This is because, in the process of singing, different singers have different faces, hairstyles, and postures. If there is a certain movement range (even sometimes there will be intense dance movements) to perform, it is impossible for the face to keep a positive and stable posture for detection at any time, and the detection efficiency is greatly reduced. Moreover, the relative position and size of the lip area will change with factors such as speech and expression. Therefore, this method is only suitable for preliminary and rough determination of lip area, which reduces some calculation work, but further accurate detection and confirmation are needed.

2.2.2. Color-Based Detection. According to the color distribution of the human face, we can find that the color of lips of most people is quite different from that of other areas of their face. Therefore, we can analyze the color intensity and color interval of lips according to different colors so as to correctly detect the parts of lips.

Because the red part of the face and lips account for different proportions. We usually choose to exclude the red gamut in the face by the red exclusion method, which is very classic and can effectively promote the solution of the problem. After reducing the influence of the red part on color detection, the situation represented by the original red area is replaced by blue and green. Here, we use the concept and meaning of three primary colors (RGB). This method can save the steps of color space conversion and tedious workload. There is also a method called the pseudocolor method, which is used to enhance the contrast between skin color and lip color. Finally, the image is converted to YIQ space, excluding the Y component, and good color discrimination can be seen. The formula is as follows:

$$\log\left(\frac{G}{B}\right) < \beta, \quad (2)$$

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (3)$$

Formula (2) is a discriminant, where G and B represent green and blue component values, respectively, which are

TABLE 1: Main parameters of corpus.

Bimodal corpus database	
Recognition primitive	Phonemes, words, sentences
Scale	Several to dozens, even hundreds of thousands
Number of people collecting corpus	From one to dozens
Visual channel information	Color or grayscale image, storage format, frame rate
Audio channel information	Background complexity, illumination, etc.

thresholds calculated by statistics. If the discriminant is satisfied, the pixel becomes a lip color pixel, otherwise it becomes a skin color pixel. Formula (3) is the formula of color intermediate conversion, which converts the image from RGB space to YIQ space, where Y is the gray component and I and Q are the chrominance components.

However, it should be noted that people have different skin tanning degrees, and each person's lips and skin color are different, so the specific colors need to be classified and discussed differently, which is not universal and brings more troubles to the specific recognition work. In addition, because singers will wear certain makeup because of social etiquette, work, and other factors, all kinds of lip gloss and cosmetics will change the color of their face, which greatly reduces the accuracy based on color detection.

2.2.3. Model-Based Detection. This method mainly uses the key points above the lip shape in the image to confirm the contour of the lip. Using the ACM algorithm, we can quickly build a matching lip model. The calculation degree of this method is very complex. If it is applied in practice, the burden of image processing will be very large, and it is difficult to correct mistakes in time once they go wrong. The relevant formula is as follows:

$$E_{\text{Snake}}(V) = \sum_{i=1}^N E_{\text{Snake}}(i) = \sum_{i=1}^N (E_{\text{int}}(i) + E_{\text{ext}}(i)), \quad (4)$$

$$E_{\text{int}}(i) = \alpha_i \|V_i - V_{i-1}\|^2 + \beta_i \|V_{i-1} - 2V_i + V_{i+1}\|^2,$$

$$E_{\text{ext}}(i) = E_{\text{image}}(i) + E_{\text{con}}(i).$$

A Viola-Jones method for lip region detection was established [18]. In this paper, we choose to fuse the abovementioned lip detection methods, and combine the face detection method with Viola-Jones method to learn from each other's strengths. A flow chart related to lip area detection is shown in Figure 1.

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'), \quad (5)$$

$$s(x, y) = s(x, y-1) + i(x, y), \quad (6)$$

$$ii(x, y) = ii(x-1, y) + s(x, y). \quad (7)$$

In Figure 1, the Haar feature extraction module is mainly used. The value of integral image ii of any image i at any pixel

(x, y) is defined as formula (8), and formula (9) and formula (10) can be obtained by calculation.

The training for AdaBoost is as follows:

$$h_i(x) = \begin{cases} 1, & \sum_{t=1}^T a_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T a_t, \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

2.3. Image Processing Techniques. The digital image is represented by a two-dimensional array [19, 20]. Image processing technology has made more and more applications in recent years [21, 22]. Each pixel is unique, with unique plane position coordinates and values, and is the basic storage unit of digital images. In order to improve the clarity of the image, we can sharpen and enhance the image.

$$d(x, y) = f(x, y) - g(x, y). \quad (9)$$

The image is preprocessed. The color image is grayed [23]. Doing so allows you to have fewer pixels, less memory for your files, and less burden on your computer when it comes to processing images. The part formula of image grayscale is as follows:

$$\text{gray} = R * 0.3 + G * 0.59 + B * 0.11,$$

$$\text{gray} = G,$$

$$\text{gray} = (R * 76 + G * 151 + B * 28) \gg 8, \quad (10)$$

$$\text{gray} = \frac{(R * 30 + G * 0.59 + B * 11)}{100}.$$

It means far greater than that. Four gray value formulas represent four different gray methods of images.

Binarization [24] makes the image black or white as follows:

$$g(x, y) = \begin{cases} 0 & (\text{gray value is less than threshold}(T)), \\ 255 & (\text{gray value is greater than threshold}(T)). \end{cases} \quad (11)$$

In the image binarization operation, T represents the gray value of pixels on the image. The confirmation of T is mainly as follows: when the gray level is higher than the threshold pixel, it is determined to be represented by the gray level value 255; otherwise, the gray value is 0, indicating the background or exceptional object area.

Take the Canny operator as an example, as shown in Figure 2:

Common convolution kernel templates are as follows:

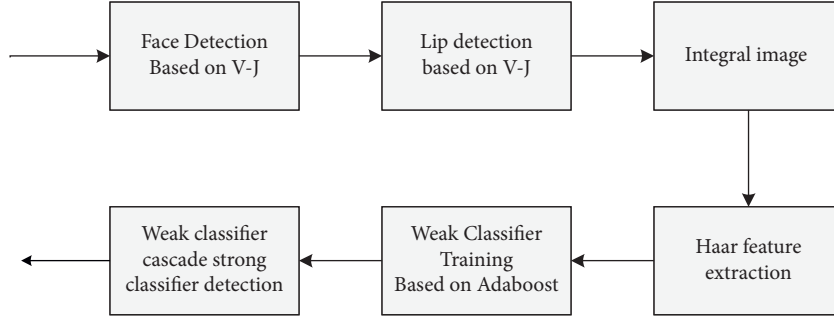


FIGURE 1: Viola-Jones lip detection flow chart. In the Haar rectangle feature.

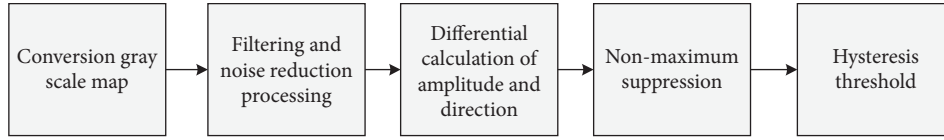


FIGURE 2: Edge detection flow of the Canny operator.

$$\begin{bmatrix} \frac{1}{16} & \frac{2}{16} & \frac{1}{16} \\ \frac{2}{16} & \frac{4}{16} & \frac{2}{16} \\ \frac{1}{16} & \frac{2}{16} & \frac{1}{16} \end{bmatrix}. \quad (12)$$

Calculate the magnitude and direction of the gradient, and the following are the x direction and y direction:

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad (13)$$

$$S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

Let the images $H(i, j)$ and C be the gradient to be calculated.

$$H(i, j) = \begin{bmatrix} A_0 & A_1 & A_2 \\ A_3 & C & A_5 \\ A_6 & A_7 & A_8 \end{bmatrix}. \quad (14)$$

We can get gradients in the x and y directions, respectively, by using the following equation:

$$\begin{aligned} G_x &= 2 \times A_5 + A_2 + A_8 - (2 \times A_3 + A_0 + A_6), \\ G_y &= 2 \times A_7 + A_6 + A_8 - (2 \times A_1 + A_0 + A_2). \end{aligned} \quad (15)$$

At this point, the gradient amplitude and direction at point C are as follows:

$$\begin{aligned} G_{C(i,j)} &= \sqrt{G_x^2 - G_y^2}, \\ \theta &= \arctan\left(\frac{G_y}{G_x}\right). \end{aligned} \quad (16)$$

2.4. Lip Feature Extraction Method. The lip feature extraction method [25] is shown in Table 2.

In this paper, we design a lip feature extraction flow chart, as shown in Figure 3, to get the final feature vector.

3. Design of a Collaborative System of Mouth Shape and Lyrics

3.1. System Development Environment. In view of the realization of the function of coordinating the subtitles in the video with the singer's mouth shape, we designed a lip-reading recognition system. The system functions are mainly divided into four functions: lip detection, feature extraction, lip-reading recognition, and automatic subtitle correction. Adjusting the appearance time of lyrics in the video makes the mouth shape coordinate with the lyrics in singing practice. The specific development environment is shown in Table 3.

3.2. Lip Reading Based on the HMM Model. The process of the hidden Markov model is double stochastic. When people speak and sing, the staying time of each lip movement is different and it is difficult to determine. This model is consistent with the process of human lip movement, can describe the pronunciation state, and is widely used in lip reading applications. In this paper, the discrete hidden Markov model (DHMM) is adopted. As shown in Figure 4, it is the composition diagram of the HMM.

Considering that pronunciation has strong continuity in time in the actual lip reading system, this paper chooses the

TABLE 2: Comparison of three extraction methods.

Method	Typical algorithm	Advantages	Disadvantages
Shape-based method	Geometric feature method, snake model method, and active contour model method	Low feature dimension; it is not easy to change	The lip movement information reflected is not comprehensive; the image has clear edges
Pixel-based method	Method based on image transformation	All pixels represent visual information together, and the loss of information is relatively small	High feature dimension; image geometric transformation sensitivity
Hybrid methods	Method based on motion analysis	Combine shape-based and pixel-based features	The algorithm is complex; it is difficult to extract the ideal contour in processing

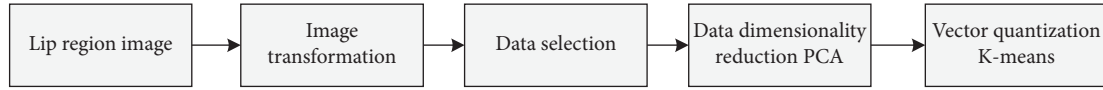


FIGURE 3: Flow chart of cascade lip feature extraction.

TABLE 3: Lip reading system development environment.

Design interface	MFC class library
Computer	Intel CORE i7 8th gen
Development tools	Visual studio 2020 integrated development environment
Operating system	Windows 10
Function library	OpenCV, MFC

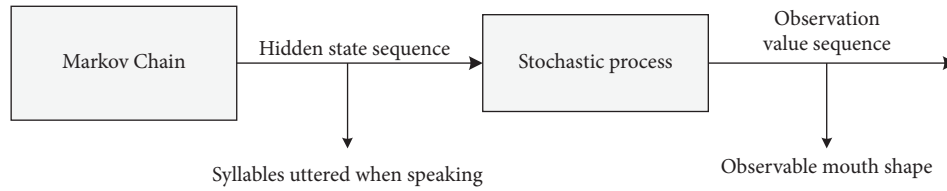


FIGURE 4: Schematic diagram of the HMM composition.

topological structure of a typical Markov chain from left to right without spanning, as shown in Figure 5.

The HMM model mainly has five parameters, which can determine the model. Reasonable selection of initialization parameters can increase the recognition rate of lip reading. Where B needs to be initialized, and N chooses a coefficient of 11.

$$\lambda = (M, N, \pi, A, B). \quad (17)$$

The values of π and A need to meet the following conditions:

$$\begin{aligned}
 &0 \leq \pi_i \leq 1, \\
 &\sum_{i=1}^N \pi_i = 1, \\
 &0 \leq a_{ij} \leq 1, \\
 &\sum_{j=1}^N a_{ij} = 1.
 \end{aligned} \quad (18)$$

$M = 4$, and the base mouth shape is shown in Figure 6.

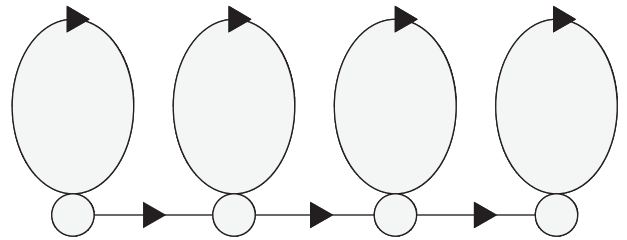


FIGURE 5: Markov chain structure.

3.3. Lip Reading Recognition and Correction. This research designs the interface module of the lip-reading system in which mouth shape and lyrics cooperate in singing practice. The whole interface strives to be concise and clear, and the functional distinction is clear. The first interface is designed as the placement area of each frame image of video interception, and all the specific processes of processing singing images are realized in this interface. The second interface button is designed as a lip area detection function. The third interface is the realization of the feature extraction function for the lip area. The fourth interface is the area where the image is finally recognized. After the recognition is successful, the data is automatically transmitted to the fifth

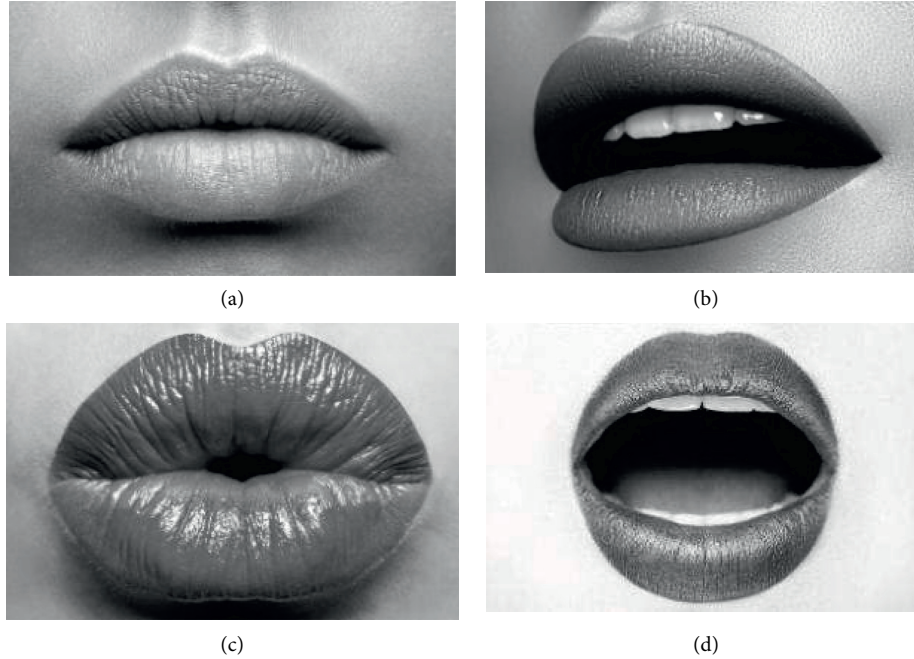


FIGURE 6: Four basic lip types of M: (a) shut up, (b) micro-opening, (c) pouting, and (d) open one's mouth.

interface to start the realization of automatic subtitle correction. At this time, the mouth shapes of the characters in the video will correspond to the subtitles, which successfully meets our experimental needs. The interface of the system is shown in Figure 7.

4. Experimental Data and Analysis

4.1. System Testing

4.1.1. UI Testing. The purpose of UI testing is to ensure that the user interface will provide users with appropriate access or browsing functions through the functions of the test objects. The content of the test is basically perfect. Check the rationality of the system operation interface designed in this paper, and the test results are shown in Table 4.

4.1.2. Performance Testing. For the experiment, the response within 2s shows that the system has the best performance; 5–10s shows that the system performance is average; if it exceeds 10s, the system performance is not good, and the system needs to be improved. The system response time is calculated as follows:

$$\text{time} = N_1 + N_2 + N_3 + N_4 + A_1 + A_3 + A_2. \quad (19)$$

The network transmission time is $N_1 + N_2 + N_3 + N_4$, the application server processing time is $A_1 + A_3$, and the database server processing time is A_2 .

We select a video whose mouth shape and lyrics subtitles are not synchronized, intercept all the images for subtitle proofreading, and test the response time of the system to deal with a large number of mouth shape images at the same

time. We tested the response time for 100, 200, 300, 400, and 500 frames, respectively, as shown in Figures 8 and 9.

From the specific case of system response time in Figure 8, we can find that the more image frames are requested to be processed, the longer the system response time is. The average processing time of each frame is 628 ms, and the system has good performance. From Figure 9, we can know that the success rate of system response is 98.80%, and the test results are satisfactory.

4.1.3. Running Time of Each Step. The running time of each system step designed in this paper is counted, as shown in Figure 10. After calculation, we can know that the image processing speed of each step is 0.36724 s, and the most time-consuming part is lip feature extraction.

4.2. Lip Area Inspection Test. The first step of the lip reading system is tested, and the images of characters from various angles are processed frame by frame for face recognition and lip region positioning. Figure 11 is a flowchart of lip region detection.

The part of the image is selected for space limitation, as shown in Figure 12. It shows face and lip recognition and detection renderings. Whether it is front or side, close-range or long-range, the system can basically accurately identify the portrait area and lip area.

4.3. Feature Extraction of the Lip Region

4.3.1. Two Cascade Methods. Using DCT-PCA and DWT-PCA, the dimension of the feature vector can be reduced by nearly 65%, and the effectiveness is greatly improved. According to the test, the recognition rate of these two

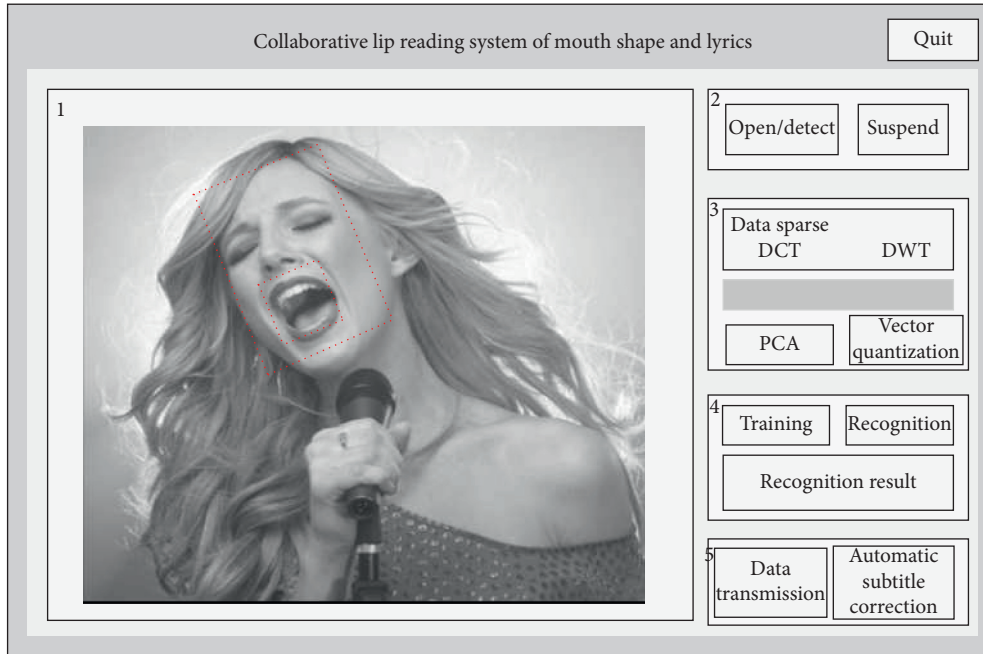


FIGURE 7: Schematic diagram of the system interface.

TABLE 4: UI testing.

Serial number	Test content	Test results
1	Interface layout	Normal
2	Text display	Normal
3	Font size	Normal
4	Garbled code	None
5	Hyperlink	Normal
6	Color style	Accord with
7	Shortcut key	Normal
8	Options button	Normal
9	Text box, dialog box	Normal
10	Clarity	Clear
11	Phenomenon of jamming or flashback	None
12	Keep a certain scale	Pass

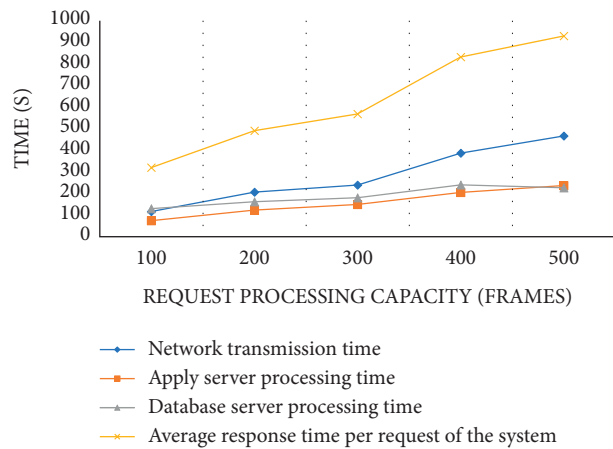


FIGURE 8: System response time.

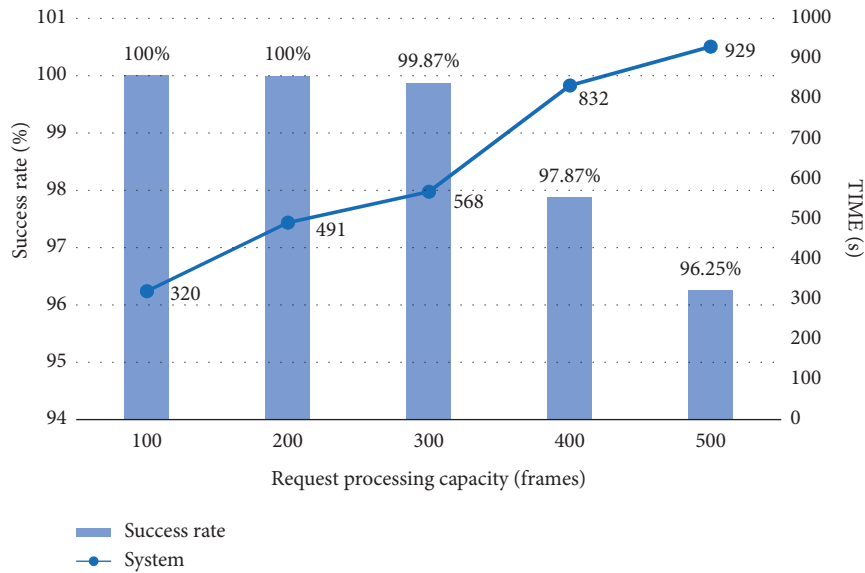


FIGURE 9: Success rate of the system response.

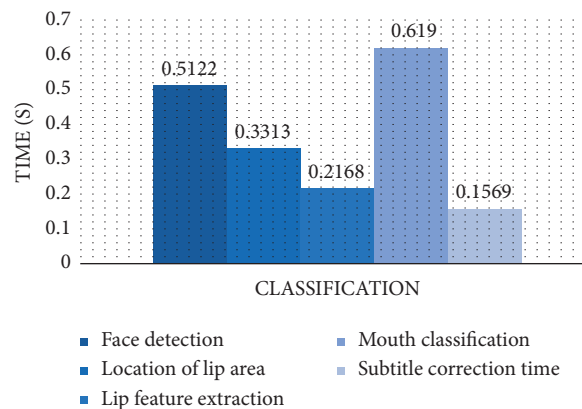


FIGURE 10: Running time statistics of each step.

methods is nearly 10% higher than that of DCT and DWT alone, as shown in Figure 13.

As shown in Figure 14, it is the change of recognition rate of two cascade methods under different feature dimensions. We can find from the figure that the higher the feature dimension, the higher the lip reading recognition rate will gradually increase within a certain range, and when it reaches a certain fixed-point peak, the recognition rate will gradually decrease slowly. This is because the number of samples in real life is limited, so when the feature dimension exceeds a certain number, it will lead to the decline of the lip reading recognition rate.

4.3.2. Labial Partial Experiment. As shown in Figure 15, the mouth shapes are classified more finely.

According to the image of human lips, lip templates are matched. Because everyone's lip shape looks different from

the standard lip shape, it needs to be optimized when identifying. As shown in Figure 16, we take the O-mouth shape of an experimenter as an example. We can see that after optimization, the O-mouth shape is obviously improved, and it is easier to classify the lip shape.

4.4. Subtitle Correction. After automatic subtitle correction, the system allows the audience to experience scoring. The specific situation is shown in Figure 17. Only the correction of 10 frames of images is shown here. It can be seen that the success rate of subtitle correction is higher than 90%, with little fluctuation and good condition. However, the audience's scoring is more subjective and fluctuates more violently.

The success rate of subtitle correction judgment is higher than 90%, which is determined by objective factors. For the audience, no matter what the subtitle correction is, the

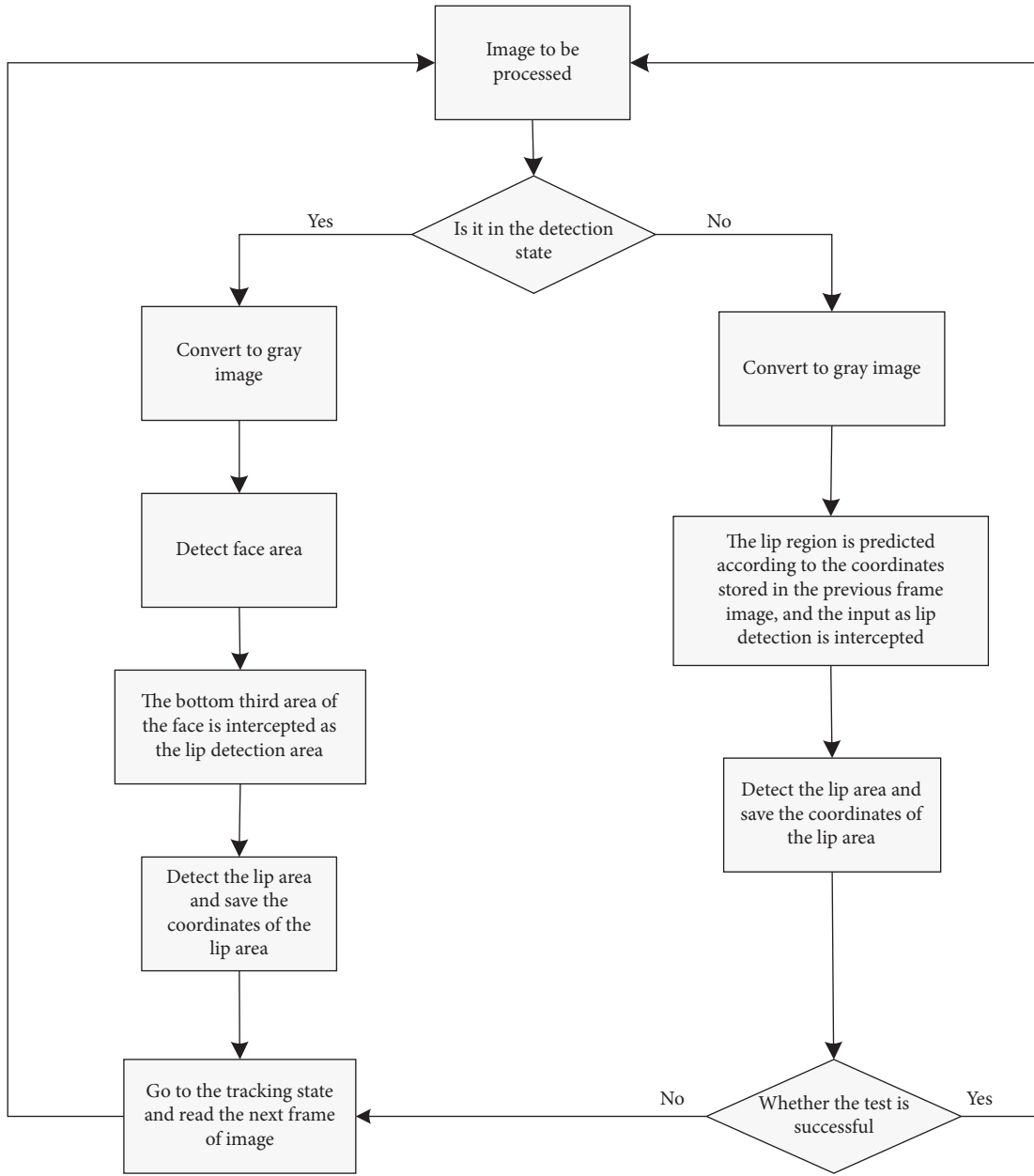


FIGURE 11: Flow chart of lip area detection.

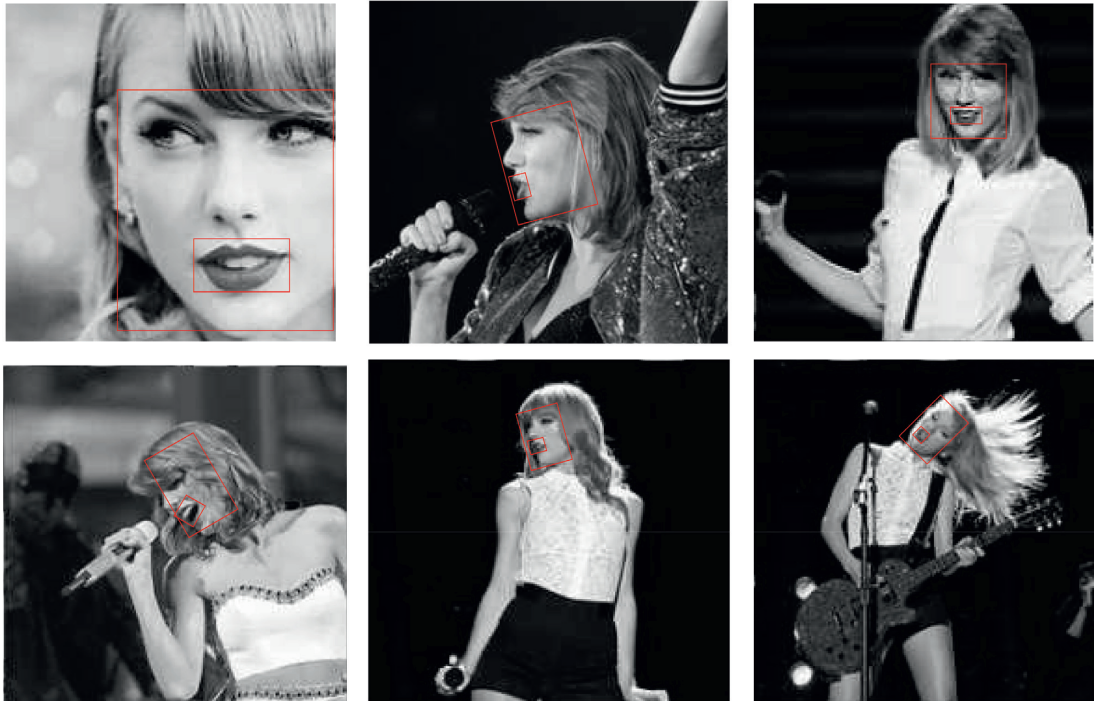


FIGURE 12: Recognition and detection renderings.

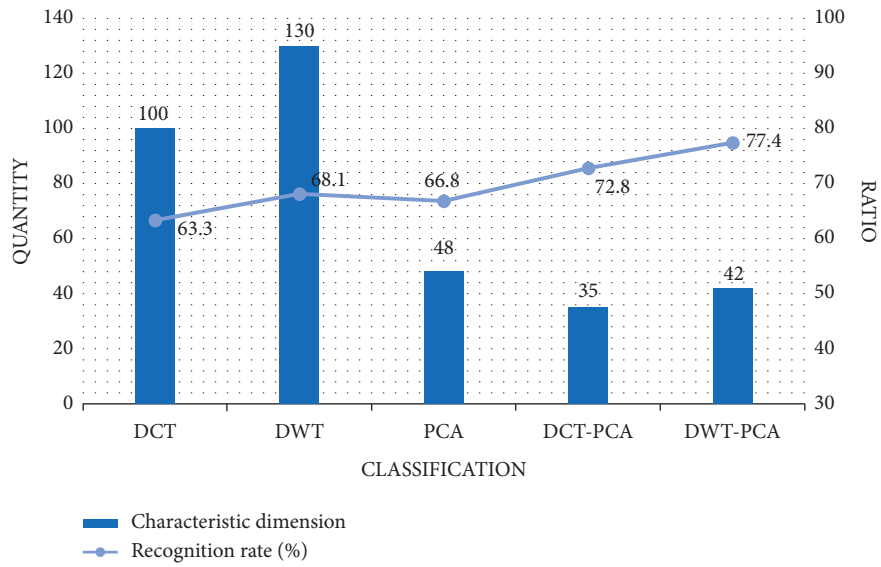


FIGURE 13: Optimal recognition rate of different features.

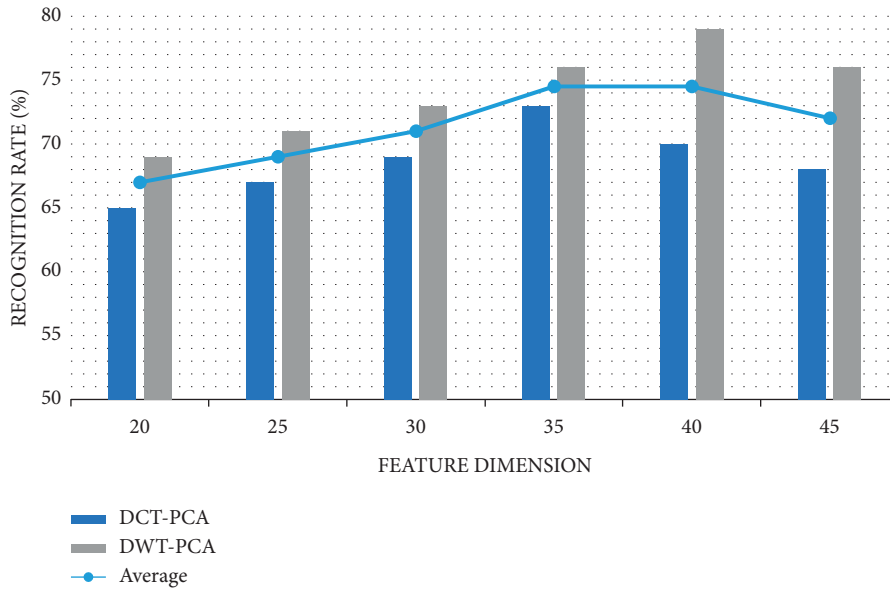


FIGURE 14: Comparison of recognition rates under different dimensions.

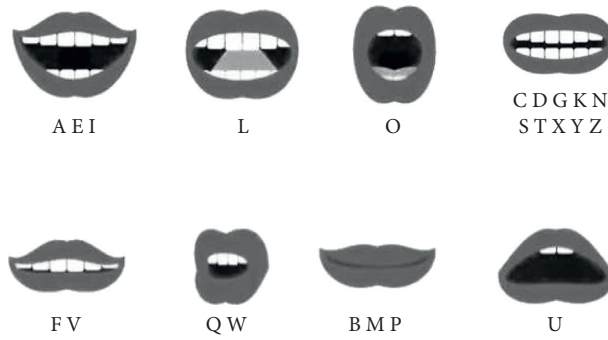


FIGURE 15: Classification according to the mouth shape of letters.

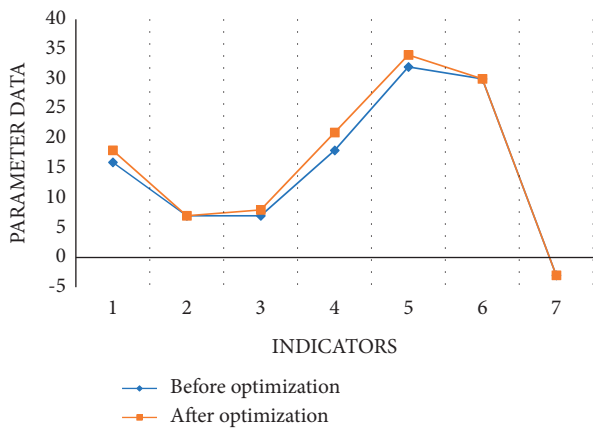


FIGURE 16: Optimization.

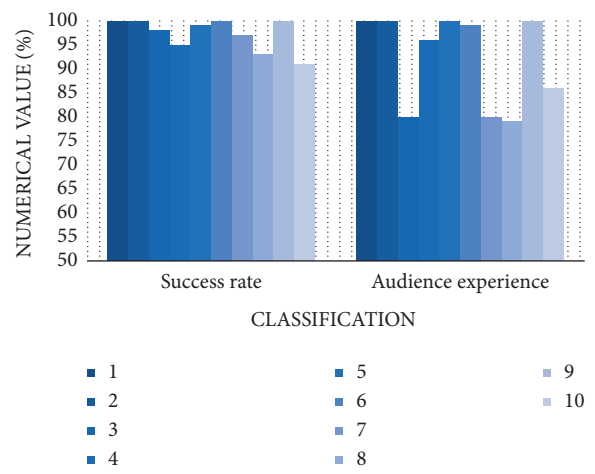


FIGURE 17: Subtitle correction diagram.

audience will subjectively judge their satisfaction with the experience. This results in a large difference between the two, which is inevitable.

5. Conclusion

Lip-reading technology is gradually moving out of the laboratory and into real life-applications, gradually integrating into people's lives without being perceived by them as its existence and role, which is its greatest success and advantage. In this paper, image processing technology is used to accurately locate the lip area for feature extraction. According to the data of the video and audio bimodal corpus, a lip reading model based on HMM is established, so that we can recognize the mouth shape accurately and adjust the lyrics synchronously with the singers in the video, thus effectively avoiding the bad experience of the audience's unsynchronized pronunciation and character in the process of watching and singing.

The research results show the following: (1) After a UI test, the functions of various operation interfaces are reasonable. The average processing time of each image frame is 628 ms. The system has good performance, and the success rate is as high as 98.80%. The test results are satisfactory. The image processing speed of each step is 0.36724 s, and the lip feature extraction part is the most time-consuming. (2) No matter the front or side, whether the character image is close-range or long-range, the system can basically accurately identify the portrait area and lip area. (3) The recognition rate of DCT-PCA and DWT-PCA is about 10% higher than that of DCT and DWT alone, and the dimension of the feature vector is about 65% lower than that of DCT and DWT alone. When the feature dimension exceeds a certain number, the recognition rate of lip reading will decrease. (4) In order to make the mouth shape of the tester closer to the standard mouth shape in recognition, some image optimization is carried out. (5) The success rate of subtitle correction is higher than 90%, and the audience's scoring is more subjective and fluctuating.

To sum up, the whole experimental results of this study ran well. The lip-reading work has achieved good results, but due to the limitations of funds, time, and technology, there are still more details to be solved in this study. For example, there are a few different types of mouth shapes, and more pronunciation and mouth shapes need to be matched in practical application; the detection of the lip area can also separate skin color and lip color more accurately; the speed of image processing by the model can be more efficient and fast [26–28].

Data Availability

The experimental data used to 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 regarding the publication of this study.

References

- [1] H. X. Yao, W. Gao, and R. Wang, "A survey of lipreading—one of visual languages," *Acta Electronica Sinica*, vol. 29, no. 2, pp. 239–246, 2001.
- [2] *Research on Lip reading Recognition Technology Based on Video Image by Tao Hong [D]*, Jiangsu University, Zhenjing, China, 2005.
- [3] *Pan Finland's Research on Intelligent Attendance System Based on Face Recognition Technology*, pp. 31–40, Zhejiang University, Zhejiang, China, 2014.
- [4] S. Tamura, K. Iwano, and S. Furui, "Multi-modal speech recognition using optical-flow analysis for lip images," *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology*, vol. 36, no. 2/3, pp. 117–124, 2004.
- [5] M. Hao, M. Mamut, and N. Yadikar, "A survey of research on lipreading technology," *IEEE Access*, vol. 8, pp. 204518–204544, 2020.
- [6] D. W. Kang, J. W. Seo, and J. S. Choi, "Monosyllable speech recognition through facial movement analysis," *Transactions of the Korean Institute of Electrical Engineers, A*, vol. 63, no. 6, pp. 813–819, 2014.
- [7] T. Saitoh, "Lip reading Technical research report of electronic information and communication society," *Emm, High Technology Enrichment of Multimedia Information*, vol. 114, pp. 29–34, 2014.
- [8] M. Kubanek, J. Bobulski, and L. Adrjanowicz, "Characteristics of the use of coupled hidden Markov models for audio-visual Polish speech recognition," *Bulletin of the Polish Academy of Sciences, Technical Sciences*, vol. 60, no. 2, pp. 307–316, 2012.
- [9] L. J. Shi, P. Feng, J. Zhao, L. R. Wang, and N. Che, "Study on dual mode fusion method of video and audio," *Applied Mechanics and Materials*, vol. 734, pp. 412–415, 2015.
- [10] I. Denis, R. Dmitry, and K. Alexey, "An experimental analysis of different approaches to audio-visual speech recognition and lip-reading," in *Proceedings of 15th International Conference on Electromechanics and Robotics "Zavalishin's Readings"*, vol. 32, no. 8, pp. 167–173, Kursk, Russia, September 2020.
- [11] L. Poomhiran, P. Meesad, and S. Nuanmeesri, "Improving the recognition performance of lip reading using the concatenated three sequence keyframe image technique," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6986–6992, 2021.
- [12] D. Hussein, D. M. Ibrahim, A. M. Sarhan, and N. M. Elshennawy, "HLR-net: a hybrid lip-reading model based on deep convolutional neural networks," *Cmc -Tech Science Press*, vol. 68, no. 2, pp. 1531–1549, 2021.
- [13] S. Isobe, S. Tamura, S. Hayamizu, M. Nose, and Y. Gotoh, "Multi-angle lipreading with angle classification-based feature extraction and its application to audio-visual speech recognition," *Future Internet*, vol. 13, no. 7, p. 182, 2021.
- [14] J. Deny, R. R. Sudharsan, and E. M. Kumaran, "An orbicularis oris, buccinator, zygomaticus, and risorius muscle contraction classification for lip-reading during speech using sEMG signals on multi-channels," *International Journal of Speech Technology*, vol. 24, no. 3, 2021.
- [15] C. Zhang and H. Zhao, "Lip reading using local-adjacent feature extractor and multi-level feature fusion," *Journal of Physics: Conference Series*, vol. 1883, no. 1, Article ID 012083, 2021.
- [16] Y. Lu and J. Yan, "Automatic lip reading using convolution neural network and bidirectional long short-term memory,"

- International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 1, pp. 2054003.1–2054003.14, 2020.
- [17] J. A. Dargham, A. Chekima, and S. Omatu, “Lip detection by the use of neural networks,” *Artificial Life and Robotics*, vol. 12, no. 1–2, pp. 301–306, 2008.
- [18] S. K. Bandyopadhyay, “Lip contour detection techniques based on front view of face,” *Journal of Global Research in Computer Science*, vol. 2, no. 5, 2011.
- [19] R. C. Gonzalez and R. E. Woods, “Digital image processing,” *Prentice Hall International*, vol. 28, no. 4, pp. 484–486, 2008.
- [20] 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, 2021.
- [21] L. Mcquillan, “Is lip-reading the secret to security,” *Biometric Technology Today*, vol. 2019, no. 6, pp. 5–7, 2019.
- [22] L. Zhang, W. Li, L. Yu, L. Sun, and X. Dong, “GmFace: An explicit function for face image representation,” *Displays*, vol. 68, no. 1, Article ID 102022, 2021.
- [23] J. C. Russ, J. R. Matey, and A. J. Mallinckrodt, “The image processing handbook,” *Computers in Physics*, vol. 8, no. 2, p. 177, 1994.
- [24] F. A. Kruse, A. B. Lefkoff, J. W. Boardman et al., “The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data,” *American Institute of Physics*, vol. 283, pp. 192–201, 1993.
- [25] Li Liu, G. Feng, and D. Beutemps, “Inner lips feature extraction based on CLNF with hybrid dynamic template for Cued Speech,” *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 1, 2017.
- [26] M. Cooke, J. Barker, S. Cunningham, and X. Shao, “An audio-visual corpus for speech perception and automatic speech recognition,” *Journal of the Acoustical Society of America*, vol. 120, no. 5, p. 2421, 2006.
- [27] I. A. Karpukhin and A. S. Konushin, “Constructing a speech audio–video corpus by aligning long segments of speech and text,” *Moscow University Computational Mathematics and Cybernetics*, vol. 41, no. 2, pp. 97–103, 2017.
- [28] N. Harte and E. Gillen, “TCD-TIMIT: an audio-visual corpus of continuous speech,” *IEEE Transactions on Multimedia*, vol. 17, no. 5, pp. 603–615, 2015.

Research Article

Construction of a Hybrid Teaching Model System Based on Promoting Deep Learning

Yingjie Sun,¹ Yang Li,¹ Yuan Tian,¹ and Wanqiang Qi ²

¹Aviation University of Air Force, Changchun, Jilin 130022, China

²School of Automotive Engineering, Jilin Teachers Institute of Engineering and Technology, Changchun, Jilin 130022, China

Correspondence should be addressed to Wanqiang Qi; qwqwall@jlenu.edu.cn

Received 22 November 2021; Revised 6 December 2021; Accepted 28 December 2021; Published 28 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yingjie Sun 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 rapid development of computer network technology in today's society has also brought new opportunities for the transformation of the teaching management mode of colleges and universities. Nowadays, most colleges and universities are gradually improving the construction of campus informatization projects, further integrating school teaching resources, centralizing the work related to the work process, and carrying out platform management to improve the daily work efficiency of various departments of the school, which is also prominent from the user experience. Humanized management out of school. This paper designs and develops a set of online teaching management system suitable for colleges and universities, realizes the informatization of college teaching management, and solves the traditional teaching methods, such as the inability of offline correspondence students to take classes in time and the limited teaching venues. Through comparison and analysis, in order to ensure that the system has better portability, this article will mainly adopt the MVC architecture model to develop and design the entire school hybrid teaching management system. The MVC architecture model is mainly developed based on the abstract design model. Strong reusability, combined with the system architecture, the main choice for database selection is development based on Oracle database technology. In the specific development process, the detailed business requirements and functional requirements of the entire school online teaching management system will be analyzed in detail. Based on this, by adopting an object-oriented approach, the system management included in the school's hybrid teaching management system will be further analyzed. Detailed analysis and design of the specific functional structure, work flow, work sequence diagram and class diagram of the online learning management and article management functional modules, and finally complete the implementation and testing of the various functional modules of the system.

1. Introduction

At present, computer technology and Internet technology have been integrated into all corners of society, greatly affecting and changing our production and lifestyles, and providing people with great convenience. From the time when the Internet was underdeveloped to the present, the convenience of Internet technology has affected all walks of life. Among them, education is a national priority and key development business, and informatization has also been applied early in education. Educational information The concept of chemistry has also begun to be widely recognized and valued. In recent years, with the continuous improvement of my country's education level, some schools have

begun to use information management systems in order to provide their own management and teaching levels. According to their own teaching management characteristics, they have designed and developed a series of education information management platforms to promote teaching Management work is developing in the direction of informatization. The emergence of a networked hybrid teaching management system has brought major reforms to the hybrid teaching management, informatization construction, sharing and release of teaching resources, networked educational administration and comprehensive management of student information in colleges and universities, and many colleges and universities have burst out The new teaching management ideas are all benefited from the rapid

development of networking and informatization. As a brand-new management method, the hybrid teaching management system is an important application of computer technology in the field of education. It strengthens the comprehensive management of schools, improves the efficiency of teachers, and advances the process of university management. At present, from a technical perspective, the current teaching management system used by private training institutions is mainly based on the C/S architecture. This architecture is simple in structure and easy to develop. The internal staff of the training institution can complete the development or commission small software. The development work is done on behalf of the development, generally the cost is low, and it is widely used in small training institutions. For colleges and universities, C/S architecture is usually difficult to meet their needs, and they mostly use teaching management systems based on B/S architecture. The B/S architecture system is more difficult to develop than the C/S architecture, but it has more powerful functions and is more suitable for school use scenarios. The B/S structure system is easy to maintain, which is conducive to the expansion of the school. On the other hand, the B/S architecture adopts a modular system architecture, and the internal modules of the system are loosely coupled, which facilitates the expansion of the department and the increase of system functions, and is more suitable for changes in the teaching management business of colleges and universities. According to the school's mixed teaching management business needs and the needs of information construction, it is very necessary to design and develop an information management system suitable for the school's own development. Through the design and implementation of this system, it can not only improve daily work and student management Efficiency, but also can reduce costs, improve their own competitiveness, and better provide students with a good learning environment and conditions.

2. Related Theoretical Methods

2.1. MVC Architecture. The whole university hybrid teaching management system mainly adopts the object-oriented design method for development and design. Through comparison and analysis, the system architecture design selection is based on the MVC (Model View Controller) architecture method. Through the use of MVC The architecture system can effectively improve the scalability and maintainability of the system. At the same time, the code reusability can be further improved in the system development process, which is convenient to reduce the system development cycle and the development cost can also be reduced. The specific situation of the relationship between the layers of the entire MVC architecture is shown in Figure 1.

The entire MVC architecture is mainly composed of three parts: view, model, and controller. Among them, the view part is mainly able to provide system users with a human-computer interaction interface, which can collect user specific operations, and can also display the content processed through the model layer. Normally, the view part

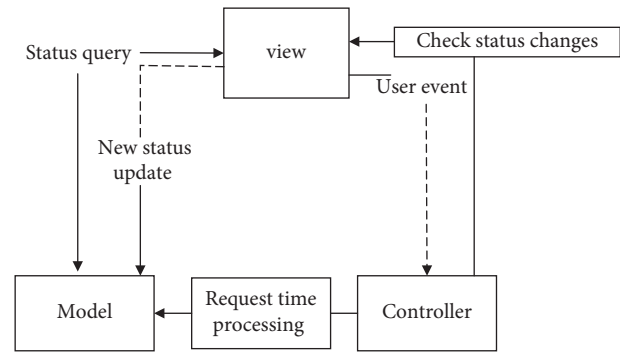


FIGURE 1: MVC architecture diagram.

is usually made of HTML elements. With the continuous development of computer technology, this module can be used to query the status of the model part and check the status change of the controller module. With the continuous development of computer technology, in order to be able to continuously enrich the machine interaction operation, provide users with more convenient and efficient interaction Experience, XML and Flash and many other elements are being applied. The model part is used to process various different types of data according to certain logic. This module can be used to query the status of the view part. In order to effectively improve the processing capabilities of various types of data, usually one model can be combined with multiple view blocks. Correspondence and processing. The controller part is the most core part of the entire MVC architecture. It can coordinate the information interaction between the view part and the model part, and process the user time information fed back by the view module to ensure the stable and reliable operation of the system. Combined with the MVC architecture diagram, it can be seen that after the user sends an operation request by using the model module in the MVC architecture, the system can call the user's specific operation request by using the view module, and at the same time feed back user event information to the controller. The controller will further retrieve the processing task data in the model module, and display the feedback and specific content information retrieved to the system user through the view module [1–6] It can greatly satisfy the effective management of the current university teaching system.

2.2. Genetic Algorithm

2.2.1. Basic Mathematical Model. The basic genetic algorithm consists of 8-tuples in the following formula, namely:

$$SGA = (C, E, P_0, M, \Phi, \tau, \psi, T). \quad (1)$$

In formula (1):

C—represents the individual coding method (binary), which is the basic factor in genetic algorithm. A combination formed by the connection of genes and chromosomes is a solution for curriculum arrangement;

E— represents the fitness evaluation function of the individual. Use this function value to distinguish whether

the individual meets the requirements; P_0 —represents the initial group, a combination of courses randomly formed by genetic algorithm, and subsequent optimization of genetic algorithm will be carried out around the initial group;

M —represents the size of the group;

Φ —represents the selection operator, and the calculation method is filtered by the fitness value. Generally, the larger the value, the easier it is to be selected;

Γ —represents the crossover operator. The new individual is generated by the random exchange of genes from the parent individual, and the single point crossover method is adopted;

ψ —represents the mutation operator, the parent individual randomly changes the individual's own gene value, and adopts a single point mutation of the gene;

T —represents the termination condition of the algorithm. When performing genetic iteration, we generally set a certain parameter value, and the operation is terminated when the parameter is iterated algebraically [7].

2.2.2. Operation Process. When using genetic algorithms to develop a hybrid teaching system, follow the steps below:

- (1) Choose an appropriate way to encode system parameters, and convert the relatively optimized solution set into chromosomes for representation;
- (2) Determine the initial group of the system;
- (3) Define the fitness function;
- (4) Genetic operation of the design system;
- (5) Control the setting of parameters such as the size of the population and the probability of genetic manipulation;
- (6) Judge whether the generated group meets a specified index, if it is satisfied, exit the calculation, if it is not satisfied, return to the operation of calculating the fitness function value [8]. The specific flow chart is shown in Figure 2.

3. System Design of Mixed Teaching Mode

3.1. System Architecture Design. In order to facilitate the system to have stronger scalability and maintainability, in the process of architecture selection, the MVC architecture is finally selected as the system architecture of the college online teaching management system, combined with specific design, the entire college online hybrid teaching management system. The detailed design of the overall architecture is shown in Figure 3.

The overall architecture of this hybrid teaching management system is mainly composed of the presentation layer, the system service layer and the data layer. The presentation layer is mainly used to provide access interfaces for system users of the university hybrid teaching management system. Through the front-end operation interface of this layer, users can provide interface services for various business operations in the system service layer [9].

The system service layer is mainly composed of core business and business support parts. The core business part

specifically includes the operation of system management, online learning management and other business functions. At the same time, in order to provide strong support for the main core business operations, in this layer The business support part has designed user authority configuration and information query and information exchange services.

The data layer is mainly used to provide services for the retrieval and storage of all data information in the hybrid teaching management system of colleges and universities, and to provide data support for the smooth progress of various core businesses in the system service layer. The business functions to be realized mainly include a variety of data information such as online course information database and video-on-demand course information database [10, 11].

3.2. System Network Topology Design. By combining the overall architecture of the system, it can be seen that the entire university hybrid teaching management system mainly adopts the MVC three-tier architecture. In order to meet the characteristics of the scattered users of the system combined with the characteristics of the system architecture, the specific network topology structure diagram of the university hybrid teaching management system. The design is shown in Figure 4.

It can be clearly found that in the course of business operation of the college hybrid teaching management system, all users in the system can use the external network and the college intranet for system login access, and the terminal devices used are computers. The network architecture can be divided into the core area and the gatekeeper isolation area. For the core area, its firewall can effectively guarantee the security performance of the system. The main method is the connection between the load balancing server and the gatekeeper isolation area, thereby providing access to the database in this area. Manage and control servers and application servers [12–14].

3.3. System Function Module Design. Online learning management is an important module of the entire university's hybrid teaching management system. It is mainly used by teachers to create, review and publish online courses. College students can also use this function for online course inquiry and registration. Combined with actual needs, the design of the functional structure diagram of the specific online learning management module is shown in Figure 5. The entire online learning management module mainly includes three sub-functions: online course management, my course management, and online learning.

The online course management sub-module is composed of two main functions: online course creation and online course review and release. The entire course management sub-module is composed of three main functions: online course inquiry, online course registration, and my course inquiry. By combining the actual needs and target requirements of the college, students can use this module to query relevant course information and combine The training goal requires course application and registration, and at the same time, it is possible to

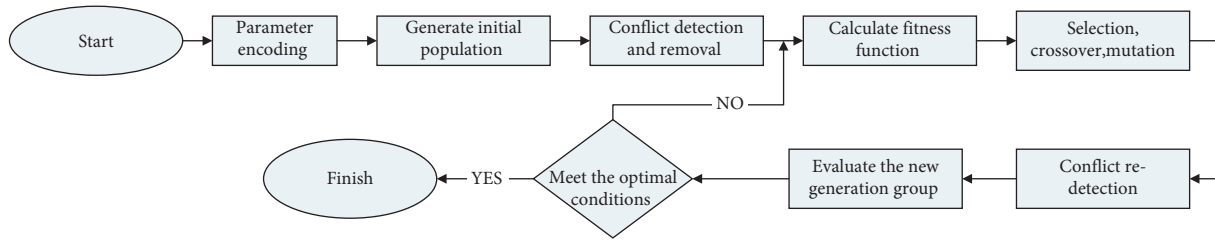


FIGURE 2: Flow chart of genetic algorithm.

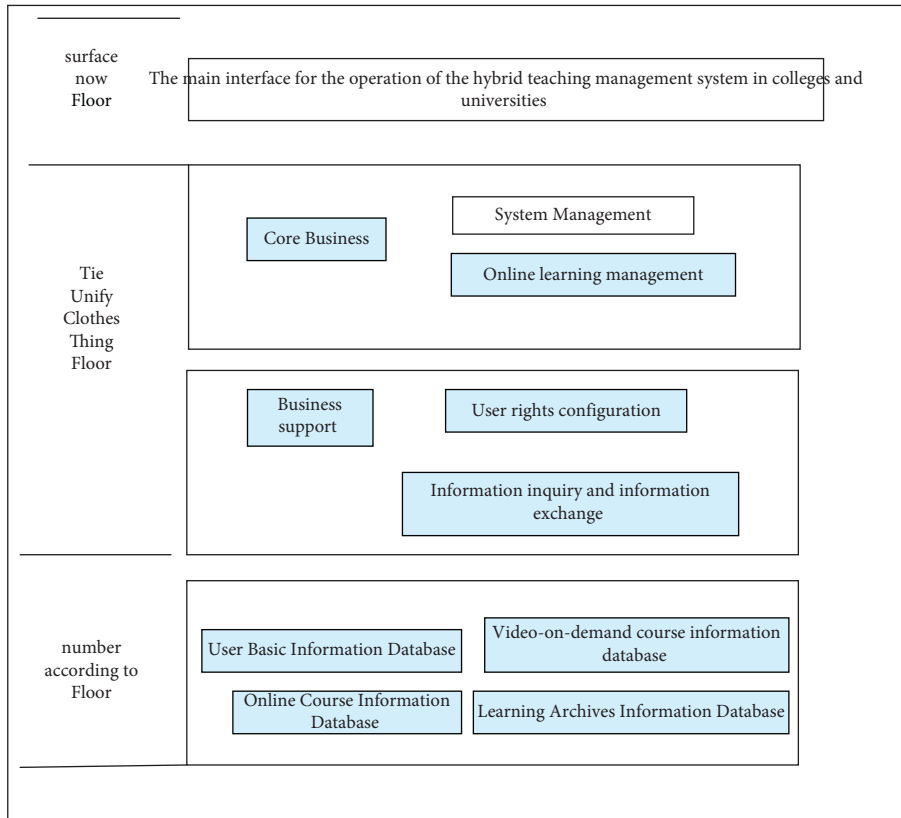


FIGURE 3: System overall architecture diagram.

inquire in detail about the course status declared by the individual. The online learning sub-module is composed of three main functions: course study material management, course video on demand, and online interaction. In combination with the online learning function requirements, the detailed online learning data management work sequence diagram is shown in Figure 6. For the various video materials uploaded by users, the system administrator is responsible for categorizing them and sorting them according to the number of views. At the same time, they set tags for each video so that users can easily and quickly find what they need through search and other means. Video learning, in the search process, on the one hand, the system administrator can combine with the actual situation to directly query the ranking of the popular course videos, on the other hand, it can also directly search through the input of the course keywords. When students are learning on video-on-demand, they

can use the interactive module to leave a message and ask questions at any time, and the teacher in charge can answer questions through the interactive platform after seeing the student’s message [15].

Combined with the working sequence diagram of the hybrid teaching management system, it can be known that the system operator clicks “online learning management” on the operation homepage, then enters “online learning,” and then completes the specific creation of the course video materials according to actual needs, and completes the course video Upload and edit on-demand resource information. After the operation is completed, after the course video on-demand resource information database is updated, college students can search for course video-on-demand resources. After entering the required query on-demand resource keywords, the system will combine the input keywords The matching information is screened in the database, and then the

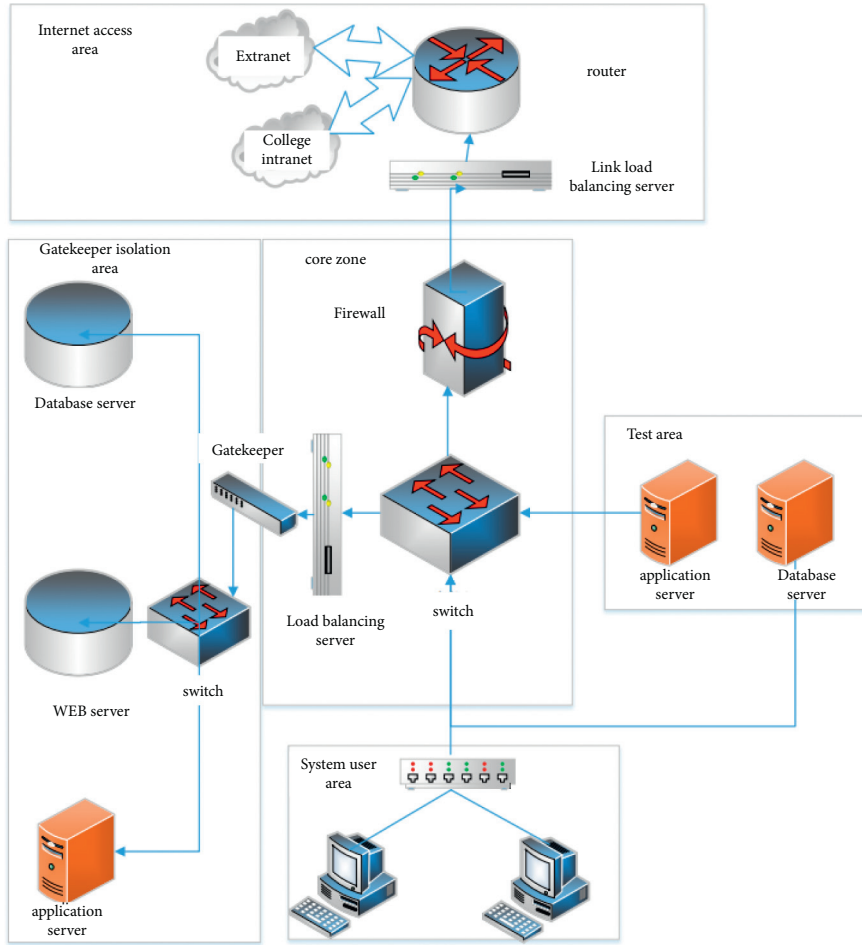


FIGURE 4: Network topology structure diagram of the hybrid teaching management system in colleges and universities.

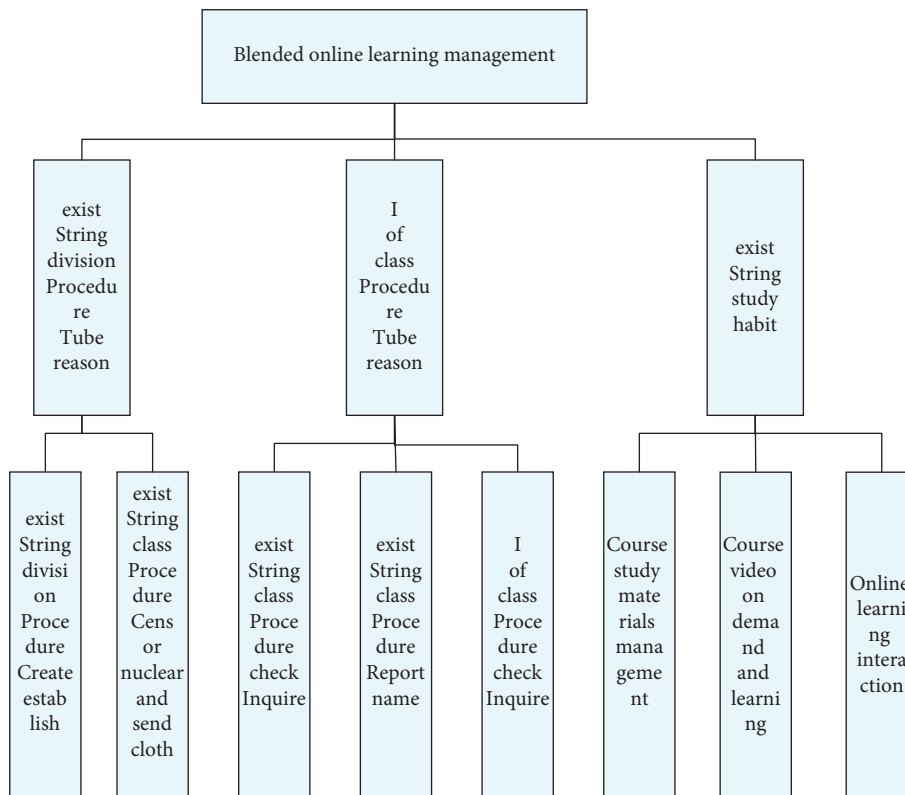


FIGURE 5: Functional structure diagram of online learning management module.

qualified resource information is displayed. In the process of watching the on-demand video, the college students can enter the interactive information of the course.

3.4. System Database Design. The database logical structure design is the design of the overall logical structure of the database. The design methods mainly include top-down, bottom-up and one-by-one Three kinds of expansion. Among them, the top-down is mainly through the design of the overall data structure, and then gradually refines until the design of the entire data structure is completed; bottom-up merges the data structure according to the logical relationship between the data, until the overall data structure is completed the design of. Combining the detailed design of the overall E_R diagram of the university hybrid teaching management system and the specific data logic structure rules, the design of the main data information tables such as online course information, test paper information and basic learning file information included in the system is explained.

For a detailed description of the online course information table, see Table 1. According to the design of the online course information table, combined with the corresponding data logical structure, this information table is composed of course name, instructor, course hours, course credits, and course duration.

4. Implementation of a Hybrid Teaching System Based on Deep Learning

4.1. System Implementation Environment. After completing the design of the functional modules and database of the hybrid teaching management system in universities, it is necessary to combine the detailed design of the system to design and implement the operation interface of each functional module of the system, combine the detailed design of the system, and control the system through the front-end operation interface. The specific implementation is fully demonstrated. In the implementation process, the corresponding hardware and software are required to support. In order to meet the stable and reliable operation of the implementation interface, the hardware and software configuration of the development side and the server side required for the system implementation are required:

- (1) Development end hardware and software configuration

On the development side, in order to facilitate the debugging of the entire college online teaching management system, in the process of selecting the terminal on the development side, the main terminal processor selected is i7-9750H, six-core CPU, memory capacity of 32GB, and hard disk type SSD. The size of the solid state drive is 1TB, and the main choice for software development is Eclipse 5.0.

- (2) Server hardware and software configuration

On the server side, combined with the requirements of system development, the server side supports RDIMM/LRDIMM, with a minimum memory of

64GB, with data protection functions such as encrypted signature firmware, secure boot, system lock, and secure erasure. Oracle is selected in the selection of database software 11g version database.

4.2. Implementation of the Mixed Teaching Management Module. Combined with the actual function design, the entire blended learning management module mainly includes three sub-functions: online course management, my course management, and online learning materials. The detailed design of the class diagram of the online learning management module is shown in Figure 7. Combined with the class diagram of the online learning management module, the entire Online Learning Management Class is mainly associated with three sub-categories: Online Course Management Class, My Cours Management Class, and Online Learning Materials Class. Among them, the Online Course Management sub-category mainly includes online Course Creation () And online Course Audit and Publication () two operation methods, used to realize the creation, review and publication of Online Course Information; My Course Management Class subclasses include online Course Query (), online Course Registration () and my Course Query () Three methods, using the My Course Management interface to implement online course query and registration and other related operations, and can be correlated with Online Course Information in the Online Course Management class; Online Learning Materials Class subclasses include course Data Management (), Course Video on Demand Learning() and online Learning Interaction() three operation methods are used to realize the management of VOD Resource Information, and use the Online Learning Interaction interface to realize the function operation of video-on-demand course information query and online interactive learning.

Combined with the specific design, the implementation of the operation interface of the online course management sub-functions in the blended learning management module is shown in Figure 8. From the online course management operation interface, it can be seen that the instructor can create the online courses he is responsible for, such as the specific courseware of the course, after-school exercises, and real test models.

The realization of the core code of the personal information management operation part in the management module of the entire hybrid teaching management system is shown in Figure 9 below:

4.3. System Test

4.3.1. Function Test. In the process of testing the online learning management module, detailed and comprehensive tests were mainly conducted on the related operations of online course management, my course management, and online learning. The specific online learning management module test case descriptions are shown in Table 2. From the online learning management function test case, it can be seen that in the process of testing the online learning

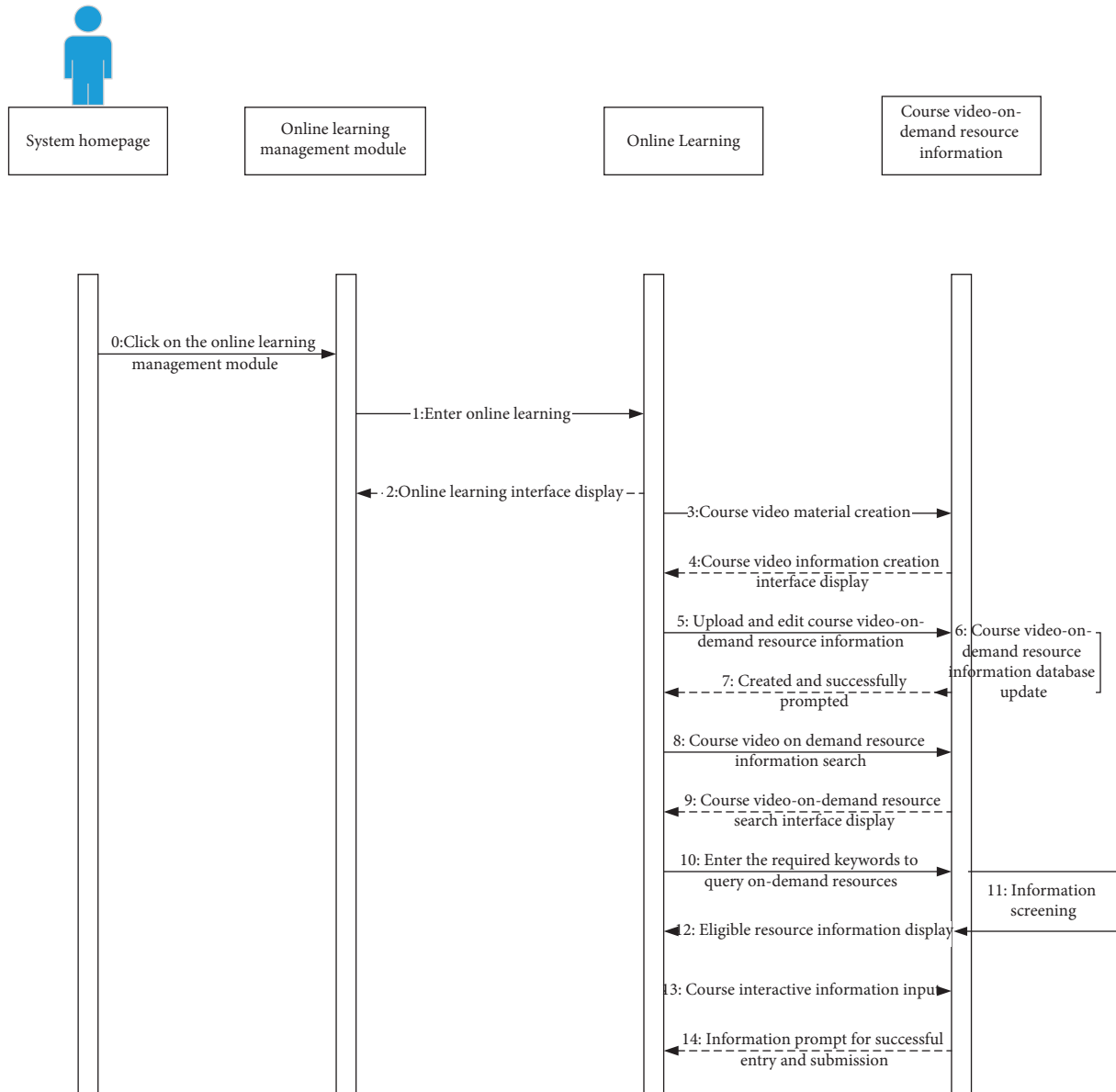


FIGURE 6: Working sequence diagram of online learning operation.

TABLE 1: Detailed description of online course information table.

Field name	Field type and length	Field explanation
Course_ID	Varchar2(20)	Course number (primary key)
Course_Name	Varchar2(50)	Course title
Course_Type	Varchar2(30)	Course type
Teacher	Varchar2(20)	Teacher
Course_Hours	Float(10)	Course hours
Chapter_Content	Varchar2(2000)	Chapter content
Course_Credit	Float(10)	Course credits
Subordinate_Specialty	Varchar2(50)	Their profession
Course_Manager	Varchar2(20)	Course manager
Exercise	Varchar2(2000)	Homework
Course_Duration	Float(10)	Course duration

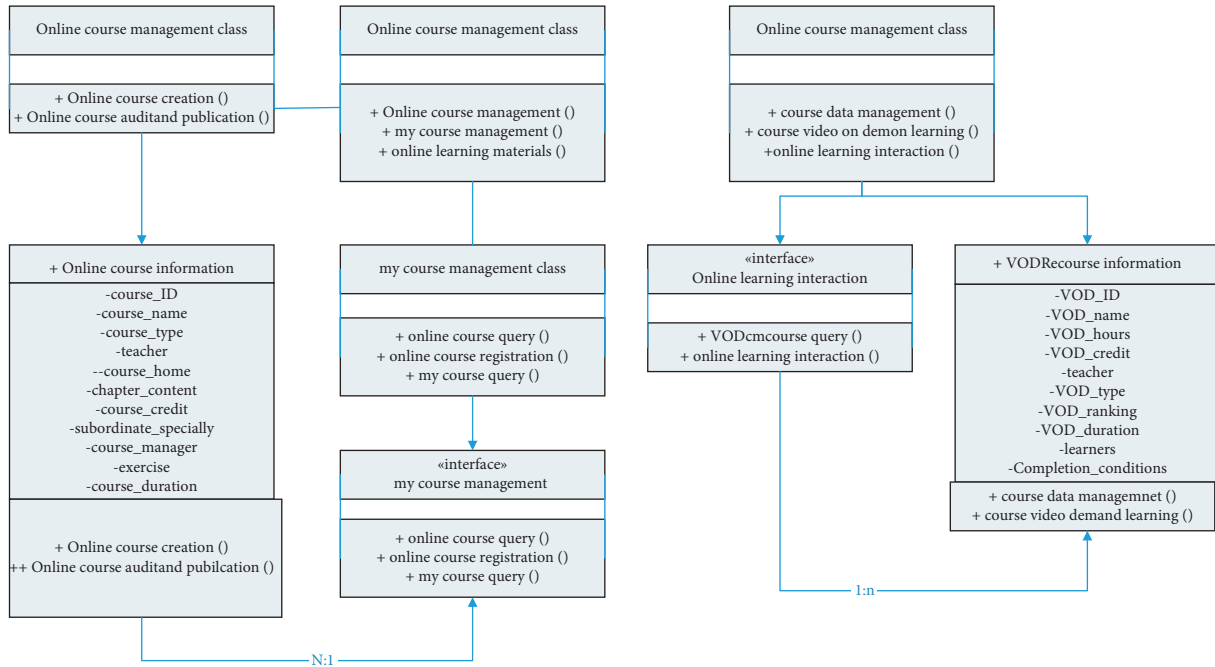


FIGURE 7: Class diagram of online learning management module.

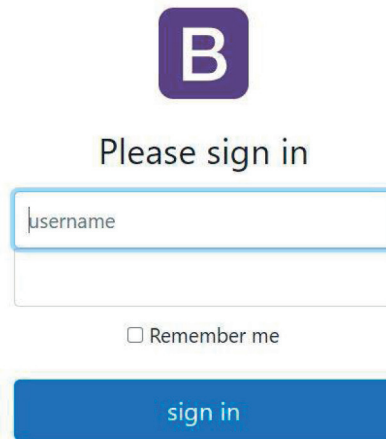


FIGURE 8: User login implementation interface.

management function, the main functions of online course management, my course management and online learning are comprehensively tested, and the actual test results can meet the expected indicators and Require.

4.3.2. *Performance Test.* In the process of testing the performance of the entire university hybrid teaching management system, the author mainly conducts the performance of concurrent users of the entire system. The detailed test situation is

shown in Figure 10. It can be seen from the concurrent performance test of the user login and processing flow that, combined with the actual performance requirements of online teaching in colleges and universities, in the specific test, the number of concurrent users selected is 300, and the main tool used in the test process is LoadRunner for testing. According to the concurrent performance test, the detailed test situation analysis table is shown in Table 3. It can be seen from the test situation analysis table that when the maximum test user is 300

```

1 public String managedept(HttpServletRequest req, HttpServletResponse resp)
2 throws Exception {
3     Admins back user = IDUtils.getBeanFromCached("back user");
4     IDUtils.modifyStr(IDUtils.getCookieTokenByReq(req), null, Admins.class);
5     List<Depts> listdepts_all =
6     IDUtils.getBeanListFromCached("cache_listdepts_all", null, Depts.class);
7     List<Depts> liststs = null;
8     String code = "0";
9     String msg_tips = "";
10    String back_page = "";
11    int count = 0;
12    int returnValue = -1;
13    if(("System administrator".equalsIgnoreCase(back_user.getAuthority()) || ("Site
14 administrator".equalsIgnoreCase(back_user.getAuthority())) {
15        liststs = IDUtils.getBeanListFromCached("cache_listdepts_all", null,
16        Depts.class);
17    }
18    if(IDUtils.isNotBlank(req.getParameter("sys_parent"))) {
19        liststs = QueryUtils.getDeptSonsById_Incache(liststs,
20        req.getParameter("sys_parent"));
21    }
22    if(IDUtils.isNotBlank(req.getParameter("keyword"))) {
23        liststs = QueryUtils.getDeptSonsByName_Incache(liststs,
24        req.getParameter("keyword"));
25    }
26    if((null == liststs) || (0 == liststs.size())) {
27        count = 0;
28        msg_tips = "No data";
29    } else {
30        count = liststs.size();
31        returnValue = 1;
32    }
33    resp.getWriter().write(IDUtils.getBack_json_list(code,
34    msg_tips,
35    back_page,
36    returnValue,
37    count,
38    IDUtils.getCurPageCon(liststs, req.getParameter("page"),
39    req.getParameter("limit")),
40    Depts.class));
41    String fromtable = table_cname;
42    String logcon = fromtable + "Search/list" + ", Operator information(" + "username: " +
43    back_user.getName() + ", User rights: " + back_user.getAuthority() + ", actual name: " +
44    back_user.getTname() + ", "+ "quit" + "Search/list" + "of" + fromtable + "information
45    (" + "Total search" + fromtable + "information: " + count + "strip");
46    String fromid = "";
47    String logtype = fromtable + "operate";
48    String adminid = back_user.getId();
49    String adminname = back_user.getName();
50    String adminname = back_user.getTname();
51    new UserLogThread(IDUtils.uuid(), adminid, adminname,
52    adminname, fromtable, fromid,
53    logcon, logtype, IDUtils.getIpAddr(req), IDUtils.getDateTime()).run();
54    return null;
55 }

```

FIGURE 9: The core code of the management module of the hybrid teaching management system.

TABLE 2: Online learning management function test case description table.

Functional test number	Testing requirements	Prerequisites	Detailed operation	Expected outcome	Results of testing
GN-003	Able to realize online course creation and online course review and release operations in the online course management sub-module	Use the identity of the instructor in school Click "Online Learning Management" in the main interface of the online teaching management system to enter smoothly	On the main interface of the displayed system, select "Online Learning Management Selection" and "Online Course Management"	Successfully enter the "Online Learning Management" page, select "Online Course Management," select the corresponding course template, and click Create. After the specific content of the online course is successfully created and imported, the database information can be updated, and prompts will be given. Wait approved, the specific content of the course can be	Pass through

TABLE 2: Continued.

Functional test number	Testing requirements	Prerequisites	Detailed operation	Expected outcome	Results of testing
GN-004	Able to realize online course query and registration in the sub-module of my course management	Use the status of a college student to click "online learning management" in the main interface of the school's online teaching management system to enter it smoothly	On the main interface of the displayed system, select "Online Learning Management Selection" and "Then My Course Management"	successfully released, and can be inquired on the system homepage Successfully enter the "Online Learning Management" interface, select "My Course Management," you can query the published online courses according to the release date or course name, and can declare the course as a student. After the registration is successful, the system can prompt to sign up success	Pass through
GN-0 05	Able to realize the video-on-demand and learning operation of the course in the online learning sub-module	Use the status of a college student to click on "Online Learning Management" in the main interface of the online teaching management system to enter it smoothly	On the main interface of the displayed system, select "Online Learning Management Selection" and "Then Online Learning"	Successfully enter the "online learning management" interface, select "online learning," you can query the selected course video in the displayed interface, and can order the required learning video according to the date and chapter, and the system can play the selected video normally, during the learning process, be able to mark the videos you watched and make corresponding notes	Pass through

Analysis Summary

Scenario Name: Scenario1
Results in Session: C:\software\weaver\flowwork\res\res.lrr
Duration: 23 minutes and 44 seconds.

Statistics Summary

Maximum Running Users:	200
Total Throughput (bytes):	72,168,112,552
Average Throughput (bytes/second):	50,644,290
Total Hits:	4,388,939
Average Hits per Second:	3,079.957

[View HTTP Responses Summary](#)

You can define SLA data using the [SLA configuration wizard](#)
 You can analyze transaction behavior using the [Analyze Transaction mechanism](#)

Transaction Summary

Transactions: Total Passed: 36,840 Total Failed: 0 Total Stopped: 0

Transaction Name	SLA Status	Minimum	Average	Maximum	Std. Deviation	98 Percent	Pass	Fail	Stop
	⊙	0.008	1.44	5.528	0.765	2.429	12,280	0	0
	⊙	0.001	0.917	9.945	1.444	2.958	12,280	0	0
	⊙	0.005	0.481	4.941	0.406	0.93	12,280	0	0

Service Level Agreement Legend: ✔ Pass ✘ Fail ⊙ No Data

HTTP Responses Summary

HTTP Responses	Total	Per-second
HTTP 200	4,061,826	2,850.404
HTTP 302	310,738	218.062
HTTP 404	16,375	11.491

FIGURE 10: System concurrent performance test.

TABLE 3: Concurrent user performance test table.

Number of test users	Expected test value (seconds)	Actual test value (seconds)
50	2.0	1.444
100	3.0	2.688
150	4.0	3.676
200	5.0	4.941

people, the actual test response time is 4.941 seconds, and the expected performance index can be met within the expected response time.

5. Conclusion

Based on the current situation of mixed teaching management, this paper attempts to design and develop a set of mixed teaching management system suitable for colleges and universities, realizes the informationization of college daily management, and solves the problems of low efficiency and high error rate of traditional manual management mode. By using the MVC architecture and Oracle database and other major technologies, the author has mainly completed the following aspects: [16–24].

- (1) A detailed analysis of the business requirements for the application of the hybrid teaching management system in colleges and universities is carried out, and the business functions required by the hybrid teaching management system in colleges and universities as well as the specific functional requirements and performance requirements of the system are proposed.
- (2) According to the specific design and implementation, the system function test and system performance test are completed by equivalence class division and boundary value analysis methods, and the test results can meet the expected index requirements.

The college hybrid teaching management system designed in this article still needs to be improved. For example, although the online learning management module already has the interactive function of online learning, it only realizes simple text communication. It is not possible to directly communicate with the instructor for the time being. Group discussions, etc., still need to be further improved.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work received no funding.

References

- [1] F. Garcia and B. Grabot, “Enterprise Web Portals for Supply CHain Coordination: A Case Study,” in *Proceedings of the IFIP International Conference on Advances in Production Management Systems*, Tokyo, Japan, September 2015.
- [2] Z. Hou, J. Gu, Y. Wang, and T. Zhao, “An autonomic monitoring framework of web service-enabled application software for the hybrid distributed HPC infrastructure,” in *Proceedings of the International Conference on Computer Science*, Kochi, India, November 2016.
- [3] B. Chen, J. Yao, and P. Wang, “A Design and Implementation of the Online Curriculum System Based on the Browser/Server Architecture,” in *Proceedings of the International Conference on Applied Science & Engineering Innovation*, Jinan, China, August 2015.
- [4] C. C. Gotlieb, *The construction of class-teacher time-tables*, IFIP Congress, Laxenburg, Austria, 1962.
- [5] M. R. Garey and D. S. Johnson, *Compute and Intractability: A Guide to the Theory of NP Completeness*, W. H. Freeman, San Francisco, 1975.
- [6] E. A. Rudnicka, “Ethics in an operations management course,” *Science and Engineering Ethics*, vol. 11, no. 4, pp. 645–654, 2015.
- [7] S. A. Roman and J. Tyler, Eds., *Mastering Enterprise Java Beans*, John Wiley & Sons, Hoboken, NY, USA, 2nd edition, 2017.
- [8] Z. K. King, *SQL Tips and Techniques*, History of China Water Resources and Hydropower Press, Beijing, China, 2016.
- [9] L. Zhou, J. Feng, and X. Meng, *Principles of Oracle Database-Design and Implementation*, Tsinghua University Press, Beijing, China, 2017.
- [10] K. Han, *Translation of Java Object-Oriented Programming Guide*, University of Electronic Science and Technology Press, Xi’an, China, 2018.
- [11] M. Zhang, *Discussion and Research on Web Application Service Architecture Based on Java Technology*, Tongji University, Shanghai, China, 2018.
- [12] W. Zhang, “Research and design of online teaching assistant system based on campus network,” *Information Technology and Information Technology*, vol. 6, no. 7, pp. 11–12, 2018.
- [13] W. Reilly, R. Wolfe, and M. K. Smith, “MIT’s CWSpace project: packaging metadata for archiving educational content in DSpace,” *International Journal on Digital Libraries*, vol. 6, no. 5, pp. 69–70, 2013.
- [14] S. Wu, “Design and Implementation of Computer Network Teaching Resources Integration System,” *Network Security Technology & Application*, vol. 6, no. 9, pp. 59–61, 2014.
- [15] J. Zhong, J. Y. Sun, N. W. Qiu, and W. Zhang, “Research and practice of the optimization and integration of experimental teaching resources,” *Laboratory Science*, vol. 34, no. 8, pp. 44–47, 2013.
- [16] A. Armoni, “A framework for the management of information security risks,” *BT Technology Journal*, vol. 25, 2017.
- [17] Y. Lin, H. Y. Su, and S. Chien, “A knowledge-enabled procedure for customer relationship management,” *Industrial Marketing Management*, vol. 35, no. 4, pp. 32–40, 2016.
- [18] P. A. Dabholkar, M. Willemijn, V. Dolen, and K. D. Ruyter, “A dual-sequence framework for B2C relationship formation: moderating effects of employee communication style in online group chat,” *Psychology and Marketing*, vol. 26, no. 2, pp. 17–22, 2017.
- [19] B. Nielsen and C. Weise, *Testing Software and Systems*, Springer, Berlin, China, 2017.

- [20] T. Kristian, "Design and implementation of a confidentiality and access control solution for publish/subscribe systems," *Innovations in Systems and Software Engineering*, vol. 5, no. 3, pp. 19–22, 2017.
- [21] T. Keith, "Assessing the quality of use case descriptions," *Software Quality Journal*, vol. 9, no. 5, pp. 21–24, 2016.
- [22] B. CHen, J. Yao, and P. Wang, "A design and implementation of the online curriculum system based on the browser/server ArCHitecture," in *Proceedings of the International Conference on Applied Science&Engineering Innovation*, Jinan, China, August 2015.
- [23] L. M. Surhone, M. T. Timpledon, and S. F. Marseken, *Relational Database Management System*, Relational Software Oracle Corporation, California, CA, USA, 2010.
- [24] S. Zhang and X. CHen, "ResearCH and Practice of Information portal System for Digital Campus," in *Proceedings of the International Conference on Information Science*, Jeju-Si, Republic of korea, April 2011.

Research Article

Research on English Achievement Analysis Based on Improved CARMA Algorithm

Lin Hu 

Jilin University of Finance and Economics Jilin, Changchun 130117, China

Correspondence should be addressed to Lin Hu; hulin@jlufe.edu.cn

Received 6 December 2021; Revised 28 December 2021; Accepted 5 January 2022; Published 24 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Lin 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.

This paper uses data mining technology to analyze students' English scores. In view of the influence of many factors on students' English performance, the analysis is realized by using the association rule algorithm. The thesis analyzes and applies students' English scores based on association rules and mainly does the following work: (1) at present, the problem of the CARMA algorithm is low operating efficiency. The combination of the genetic algorithm's crossover, mutation, and the CARMA algorithm realizes the fast search of the algorithm. The simulation results show that the operation performance of the algorithm is greatly improved after the crossover and mutation operations in the genetic algorithm are applied to the CARMA algorithm. The simulation results show that the mining accuracy of the improved algorithm is 97.985%, and the mining accuracy before the improvement is 92.221%, indicating that the improved algorithm can improve the accuracy of mining. (2) By comparing the mining time of the improved CARMA algorithm, the traditional CARMA algorithm, the FP-Growth algorithm, and the Apriori algorithm, the results show that when the number is 6,500, the mining efficiency of the improved CARMA algorithm is twice that of the other three algorithms. As the amount of data increases, the effect of improving mining efficiency gradually increases. (3) By using the improved CARMA algorithm to analyze students' English performance, it is found that the quality of student performance is strongly related to the quality of daily homework, and if it is related to the teacher's gender, professional title, etc., it is recommended that schools should pay more attention to homework during the teaching process.

1. Introduction

Facing a large amount of information and data, the rapid development and wide application of data mining technology effectively solve the problem of how to use massive amounts of information. Various data mining analysis methods are also increasing. Classification, prediction, association rules, clustering, etc., are all frequently used data mining methods. Among them, association rules are widely used in big data mining as a rule-based machine learning algorithm [1]. Association rules find data with certain relevance and relevance from massive data, and describe the common attributes or laws of these data. Association rules are a simple and practical analysis technique, which originally originated from shopping basket analysis and goods placement, and increases the sales of goods by finding the connections between goods. The current association rules

have been applied in many fields, such as the financial industry and e-commerce. The Apriori algorithm, CARMA algorithm, and FP-Growth algorithm are the classic algorithms in association rules [2–4].

In recent years, data mining has also been widely used in the education field, especially in the application of student performance. Through the analysis of the academic performance, it is possible to understand the student's learning situation, find the student's weaknesses, better check the deficiencies and make up for omissions, promote the improvement of students' academic performance, and at the same time improve the school's teaching level. However, the current analysis of student performance is mainly simple performance query, deletion, etc., and a series of behavioral factors related to academic performance have not been excavated and analyzed, such as the impact of students' daily behavior on learning [5]. However, it is well known that

there is a close relationship between student learning behavior and academic performance, so the current role of data mining has not been fully utilized [6]. In the face of this situation, this article conducts a more in-depth analysis of the application of data mining in the analysis of high-efficiency students' English scores and specifically selects association rules to mine student scores. The level of students' English performance is also inextricably linked to factors such as the school's teachers and their teaching conditions. Therefore, it is suitable to use the association rule algorithm to find the relationship between English performance and student behavior, teachers, and other factors so as to propose English teaching suggestions.

2. Related Theories

2.1. Data Mining. Big data [7] refers to data sets that cannot be collected, processed, and analyzed within a tolerable time frame using traditional data processing and analysis techniques or software and hardware tools. Compared with the traditional dataset, it has 4V characteristics [8]: scale (Volume), diversity (Variety), value (Value), and effectiveness (Velocity). According to different mining tasks, data mining processing methods can be divided into classification analysis, cluster analysis, association rule mining, and sequence analysis. Blackett further divides the processing methods into 3 categories according to the degree of data analysis [9]: descriptive analysis that uses historical data to summarize, predictive analysis that uses related technologies to predict future probabilities or trends, and discovers hidden data relevance to help users make a regular analysis of decision making. Cluster analysis is used to describe data, which classifies data by measuring the similarity between data; classification analysis obtains classification rules based on training data, so it is often used for predictive analysis; the purpose of association rule mining is to find and extract valuable hidden patterns, and the idea is also applicable to sequence pattern discovery. Among them, classification belongs to supervised learning, and clustering and association rule mining belong to unsupervised learning. As commonly used machine learning algorithms, these three types of methods have received more attention and research in academia, and many methods and theories have been developed. Big data mining is specific, and the algorithm definition is shown in Figure 1. This article will expand the relevant content in sequence according to classification, clustering, and association rule mining.

2.1.1. Classification Algorithm. According to the mental framework in Figure 1, the classification algorithm of streaming data is first introduced. Classification is a supervised learning method [10]. Its purpose is to discover the classification rules through the examples in the training data set and the class labels to which the examples belong and use the classification rules to predict the class label of the unknown instance. The classification process has two stages: (1) the classification rule learning stage, which analyzes and summarizes the label situation of each instance class in the

training set, discovers the classification rules, and creates a suitable classifier model; (2) the classification stage of the test set. In this stage, the classification accuracy of the classifier model needs to be evaluated on the validation set. If the accuracy is good, the classifier is used to preclassify the instances in the test set; otherwise, the classifier model needs to be retrained.

Streaming data classification algorithms are mainly classified into two categories according to the number of classifiers included [11] as follows: single model algorithm and integrated classification algorithm. The single-model algorithm improves and expands the traditional batch classification algorithm and is equipped with a specific concept drift processing mechanism so that it can adapt to the changes in the data in the stream. According to traditional classification algorithms, single-model algorithms can be divided into several types based on decision trees, support vector machines, and Bayesian models. The ensemble learning algorithm trains N base classifiers based on data in different time periods, combines each classification model into an ensemble classifier through a certain combination method, and synthesizes the results of each base classifier according to a certain mechanism to obtain the result of the ensemble model, making the integrated model more robust and able to deal with the problem of concept drift flexibly.

2.1.2. Clustering Algorithm. Big data clustering is an unsupervised learning method [12], which does not require predefined categories and training samples. Its purpose is to divide the unknown data set into several clusters so that the data points in the same cluster are as similar as a possible degree, and the data points in different clusters have as much dissimilarity as possible. There are two main stages in the clustering process:

- (1) Define the similarity function and use it as the basis for judging whether the data are similar.
- (2) Select the appropriate clustering algorithm to divide the data objects into different clusters. Big data clustering analysis methods can be divided into single-machine clustering methods and multi-machine clustering methods [13]. The realization of the single-machine clustering method is mainly through the optimization and expansion of the classic clustering algorithm and the dimensionality reduction or sampling of the data set, while the mainstream multimachine clustering method is mainly reflected in the architecture and other aspects. There is no essential difference between the class algorithm and the classic clustering algorithm.
- (3) Association rule algorithm big data association rule mining refers to discovering hidden information relevance from a large amount of data and analyzing the relevance to help decision makers make correct decisions. The strength of association rules can be measured by support and confidence [14]: if there is a rule $X \rightarrow Y$, the support is the frequency of transactions X and Y appearing in the transaction

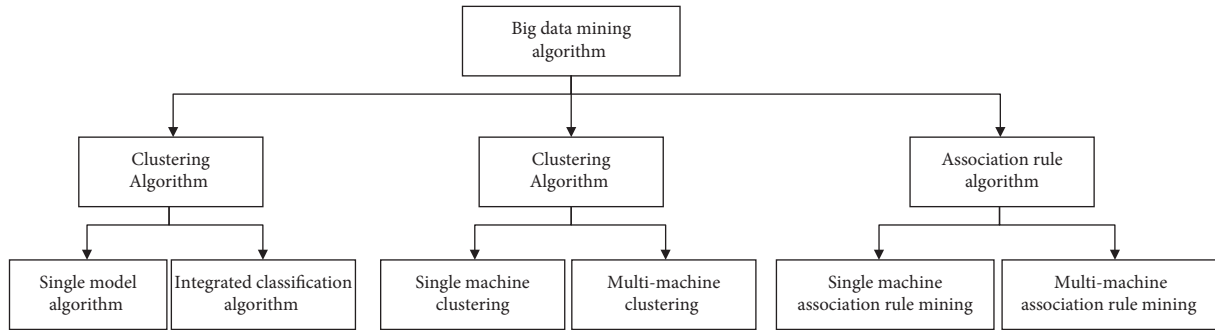


FIGURE 1: Big data mining method.

database, which is used to judge whether the rule has utility value; confidence: this degree is the frequency of the occurrence of transaction Y when transaction X appears and is used to judge the correctness of the rule. There are two main steps in association rule mining: (1) find out the rule whose support degree is not less than the minimum support degree from the data called “frequent pattern” or “frequent item set;” (2) find out the confidence degree not less than the minimum from the frequent pattern. The strong association rules with confidence thresholds focus on the mining of frequent patterns. Association rule mining can be divided into basic single-machine association rule mining and multimachine parallel association rule mining based on the distributed computing framework [15].

2.2. Association Rule Method. Association rules refer to the frequent patterns, correlations, or causal structures that exist between item sets or object sets in log data, relational data, or other information carriers. The acquisition of association rules is mainly through data mining methods to find frequent patterns from a large number of event record databases [16] and discover new patterns and potential knowledge by analyzing frequent situations.

2.2.1. Algorithm Type. (1) Apriori algorithm. The Apriori algorithm based on frequent item sets is one of the classic algorithms in the field of association rule mining [17]. The data structure of the Apriori algorithm is simple, clear, and easy to understand, but the application of the Apriori algorithm requires multiple scans of the database, which is expensive. Although the Apriori algorithm itself has made some optimizations, the time and space complexity of using this method is large, and the efficiency is not high. The specific steps are as follows:

- (1) Scan the original transactional data and extract the candidate 1 item set. According to the given minimum support s , extract the item set whose frequency of the item set is not less than the support s to form a frequent 1 item set.
- (2) Scan the original transactional data and extract the set of candidate 2 item sets, judge each 2 item sets in

the set, exclude the 2 item sets that do not contain frequent 1 item set, and form a new set as the candidate set. According to the support degree s , the frequent 2 item sets are filtered out.

- (3) Repeat the idea of the second step, get frequent 3 item sets, frequent 4 item sets,...frequent n item sets, and $n+1$ item sets do not meet the support threshold condition.
- (4) Given the confidence level, judge whether each frequent item set rule is a valid rule, calculate the lift of the rule, and judge whether the strong rule is valid.

The steps of the Apriori algorithm are shown in Figure 2.

(2) FP-Growth algorithm. The advantage of the FP-Growth algorithm is that it mines all frequent item sets without generating a large number of candidate sets. With the generation of large-scale data, the processing capabilities of data mining in a single-machine environment are challenged in both storage and computing. A parallel computing environment has become the first choice for solving big data processing problems. Some researchers have proposed parallel algorithms based on multithreading [18]. Although the pressure of storage and calculation has been relieved to a great extent, the limitation of memory resources has become the algorithm of the bottleneck of expansion; the FP-Growth algorithm is based on the Apriori principle. It finds frequent item sets by storing the data set on the FP (Frequent Pattern) tree but cannot find the association rules between the data. The FP-Growth algorithm only needs to scan the database twice, while the Apriori algorithm needs to scan the data set once when seeking each potential frequent item set. The process of the FP-Growth algorithm is to first construct the FP tree and then to use it to mining frequent item sets. When constructing the FP tree, you need to scan both sides of the data set. The first scan is used to count the frequency, and the second scan is to consider frequent item sets.

The steps of the FP-Growth algorithm are shown in Figure 3.

(3) CARMA algorithm. The CARMA algorithm was proposed by Christian Hidber and is a classic algorithm in association rules. It uses the iterative method of layer-by-layer search to find the relationship between the item sets in the database to form rules. The process consists of connection (matrix-like operations) and pruning (removing unnecessary intermediate results). The concept of item sets

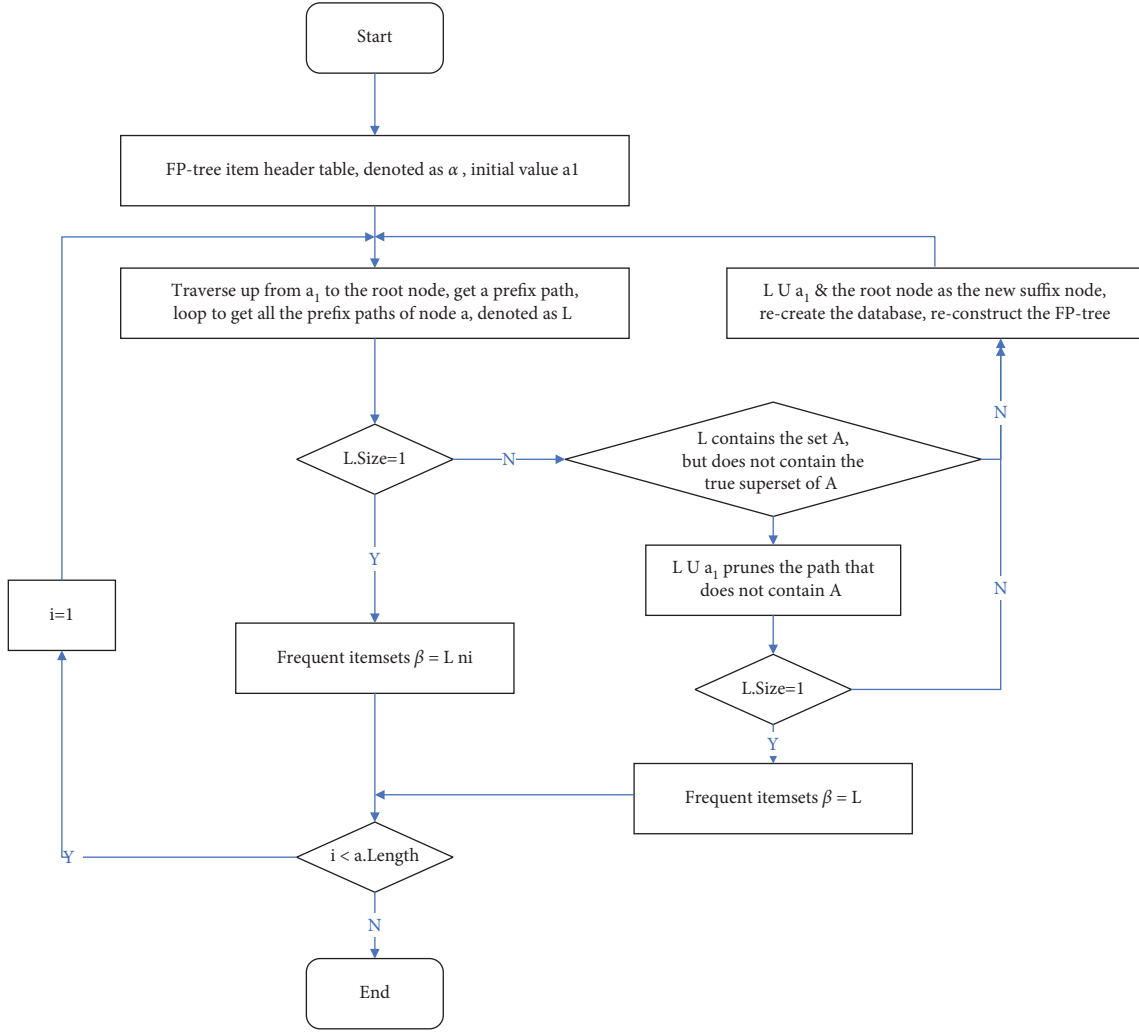


FIGURE 3: Flowchart of FP-Growth algorithm.

The calculation method of confidence is

$$\text{Confidence}(a \Rightarrow b) = p\left(\frac{b}{a}\right). \quad (1)$$

The calculation method of support is

$$\text{Support}(a \Rightarrow b) = p(a \cup b). \quad (2)$$

3. Improving the CARMA Algorithm

The CARMA algorithm is a classic association rule algorithm, which has low memory usage and continuous data processing features. However, the CARMA algorithm is not fast, and the mining efficiency is low. The way of solving this problem is one difficult question. This section proposes to introduce the genetic algorithm into the CARMA algorithm to improve the accuracy of the algorithm.

3.1. Basic Principles. The CARMA algorithm is derived from the system model. A system model refers to a model that includes inputs and outputs, including a random model and

a deterministic model. Random models can be seen everywhere in practice; this is due to the randomness of the system due to noise interference [20]. Figure 4 shows the structure of the random model.

The stochastic system can be expressed by the following formula:

$$A(z)y(k) = B(z)u(k) + \frac{D(z)}{C(z)}v(k),$$

$$y(k) = \frac{B(z)u(k)}{A(z)} + \frac{D(z)}{A(z)C(z)}v(k), \quad (3)$$

$$\text{Set } G(z) = \frac{B(z)}{A(z)},$$

$$H(z) = \frac{D(z)}{A(z)C(z)}.$$

In the definition symbol, $G(z)$ is called the system model transfer function, and $H(z)$ is called the noise model transfer function.

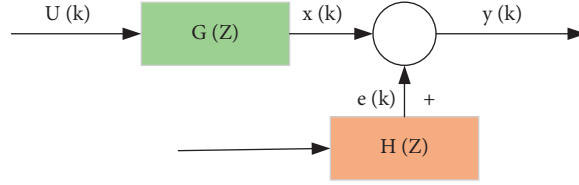


FIGURE 4: Structure diagram of random model.

When $C(z) = D(z)$, $u(k) = 0$, we get $A(z)y(k) = B(z)u(k) + v(k)$; this is called the “autoregressive model.”

When $C(z) = 1$ and $u(k) = 0$, $A(z)y(k) = B(z)u(k) + D(z)v(k)$ is obtained, which is called the “controlled CARMA model.”

The CARMA algorithm is an online algorithm, which consumes less memory during the calculation process, so it is widely used in online mining. However, the disadvantage of this algorithm is that it runs slightly slower. In order to improve the efficiency of the algorithm, this paper proposes to combine the genetic algorithm and the CARMA algorithm to form the CARMA algorithm based on the genetic algorithm to improve the efficiency of the algorithm.

3.2. Genetic Algorithm. Genetic algorithm is a bionic algorithm born out of natural selection in Darwin’s theory of evolution [21]. The algorithm flow of the process diagram is shown in Figure 5.

It can be seen from the figure that the gene is the most basic unit in the structure of the genetic algorithm, and the gene constitutes the chromosome and finally constitutes the population. A detailed analysis is given below:

- (1) **Gene:** In biology, a gene is a nucleotide sequence that produces a polypeptide chain. The DNA that contains genetic information in this sequence is the smallest unit of genetic behavior. In order to simulate gene behavior, the concept of a gene is added to the genetic algorithm [22].
- (2) **Chromosomes:** According to the principles of genetics, chromosomes are composed of a limited number of genes, and chromosomes carry a lot of information. There are two ways to encode chromosomes in genetic algorithms:

One is binary numbering, and the other is cross coding. The following two pairs of codes are analyzed. The principle of binary encoding is to convert design variables into binary encoding. Assuming that there are variables $(1, 2, \dots) p_i, i = n$, then the corresponding binary string length is set to $(1, 2, \dots) l_i, i = n$, and the decimal value $(1, 2, \dots) M_i, i = n$. The calculation method is

$$M_i = \frac{(2^{l_i} - 1)(p_i - a_i)}{b_i - a_i}. \quad (4)$$

- (3) **Population:** In order to simulate the biological population, a limited number of chromosomes are

integrated as the population number in the genetic algorithm. The population size setting directly affects the efficiency of the algorithm. The population size of the genetic algorithm designed in this paper is set in 150 pieces.

- (4) **Fitness:** a fitness function is introduced for the evaluation of indicators. The fitness calculation methods are as follows: when calculating the maximum value of the objective function, the method is calculated according to the following formula:

$$\text{Fit}(f(x)) = f(x). \quad (5)$$

The calculation method of the minimum value of the objective function is calculated according to the following formula:

$$\text{Fit}(f(x)) = -f(x). \quad (6)$$

Variant calculation method for solving the objective function optimization problem is as follows: when the minimum value needs to be calculated, there are

$$\text{Fit}(f(x)) = \begin{cases} c_{\max} - f(x) & f(x) < c_{\max} \\ 0 & f(x) \geq c_{\max} \end{cases}, \quad (7)$$

where c_{\max} represents the maximum estimated value of the function. When the objective function is the maximum, then

$$\text{Fit}(f(x)) = \begin{cases} -c_{\min} + f(x) & f(x) > c_{\min} \\ 0 & f(x) \leq c_{\min} \end{cases}, \quad (8)$$

where c_{\min} represents the minimum estimated value of the function. The objective function fraction calculation algorithm is as follows: when the minimum problem needs to be calculated, there are

$$\text{Fit}(f(x)) = \frac{1}{1 + f(x) + c} \quad c \geq 0, \quad c + f(x) \geq 0. \quad (9)$$

When the objective function calculates the biggest problem, there are

$$\text{Fit}(f(x)) = \frac{1}{1 - f(x) + c} \quad c \geq 0, \quad c - f(x) \geq 0. \quad (10)$$

The fitness function changes mainly include linear changes, power function changes, and exponential changes.

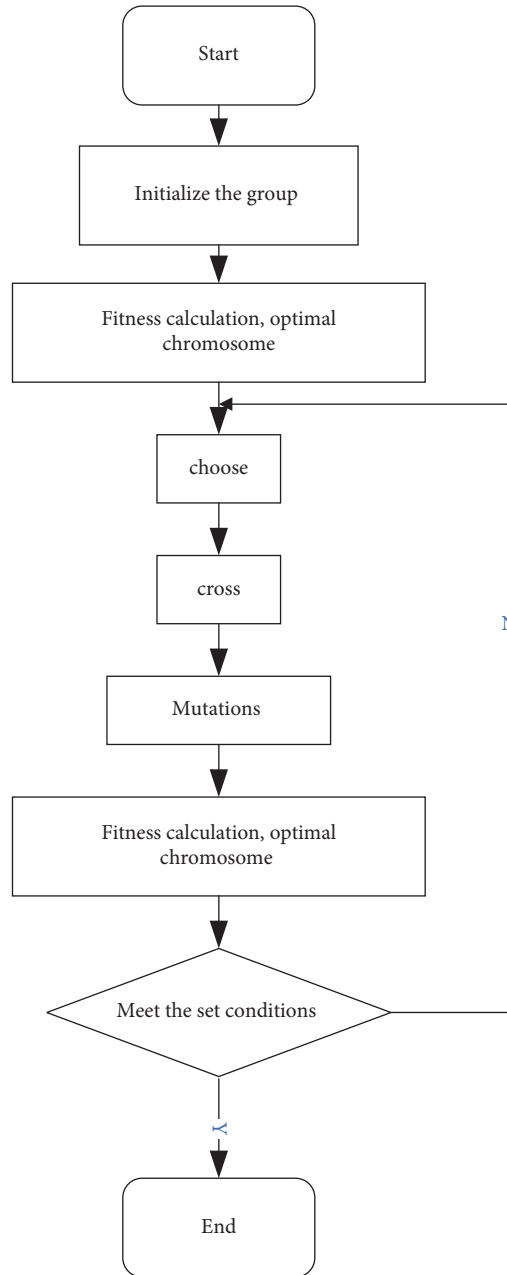


FIGURE 5: Flowchart of genetic algorithm.

(5) Selection: The genetic algorithm introduces selection to select the superior individual from the group, and the elimination of inferior individuals makes the population better. At present, the selection of calculation methods in genetic algorithms is mainly based on deviation, expectation, probability, and so on. This article chooses according to the compass calculation method. The probability calculation method is as follows:

$$P_t = \frac{f_t}{\sum_{i=1}^M f_t} \tag{11}$$

- (6) Crossover: It refers to the gene exchange between two parents, and the algorithm search ability is enhanced after crossover.
- (7) Mutation: It is easy for genes to crossover all combinations during the crossover process. In order

to prevent the partial algorithm, the optimal case introduces variation. The mutation rate in the genetic algorithm cannot be set too high. If it is too high, the algorithm will no longer have the search ability.

3.3. Improved CARMA Algorithm Model. The principle of the CARMA algorithm based on the genetic algorithm proposed in this paper is to use the genetic algorithm for frequent item search. However, the CARMA algorithm only implements mining, which reduces the operating efficiency of the CARMA algorithm. The way to improve the efficiency of the CARMA algorithm is to constrain the crossover operator and mutation algorithm of the genetic algorithm, which improves the algorithm operation efficiency. According to the CARMA algorithm idea for algorithm design [23], the designed algorithm based on a genetic algorithm, i.e., the CARMA algorithm flow, is shown in Figure 6.

From Figure 6, we can see that the improved CARMA algorithm first reads student information from the MySQL database and then divides the data in the MySQL database into several sub-data sets and stores them in the memory; second, it uses the rules in the CARMA algorithm to control the data in the memory database. Strong decomposition is carried out; the decomposed data are merged; the genetic algorithm is applied to the CARMA algorithm to achieve strength search.

From Figure 6, we can see that the core idea of the algorithm consists of two parts: (1) the MySQL database is split into multiple subdatabases, the data are transferred to the memory, and then, the operation of merging the strength set is implemented; (2) the genetic algorithm is introduced into the strength set to find the optimal strength set. The use of the genetic algorithm greatly reduces data I/O operations, thereby improving the operating efficiency of the CARMA algorithm. The fitness function setting is performed after completing the encoding according to the above operations. When genetic algorithms need to solve problems, the CARMA algorithm reads the database efficiency problem. For this reason, the fitness function constructed in this paper includes confidence and support as two variables of persistence. The calculation method of the fitness function is as follows:

$$f(x) = aS(x) + bB(x). \quad (12)$$

In the above formula, a and b are constants, $S(x)$ represents the degree of support, and $B(x)$ represents the degree of confidence. Next, the genetic algorithm selection operation is performed. As analyzed above, this article chooses the probability calculation method to choose. Algorithm 1 describes the steps of the CARMA algorithm based on the genetic algorithm in detail.

4. Experiment

4.1. Experimental Data and Operating Environment. The experimental system configuration environment is the

operating system Windows7 64-bit Professional Edition; memory is 16 GB (16 GB × 1) DDR42666 MHz; hard disk is SSD256 G; CPU frequency is 2.2 GHz, six cores/twelve threads, simulation level; the platform is MATLAB7.0.

Data source: the experimental data come from the database stored in the student status management system of a school in the past three years (the student status management system database is an open-source MySQL database). At present, the amount of data stored in the student status management system reaches 500 Mb. The first step is to divide the database. This algorithm is divided into 20 data sets. The data are read into the memory, respectively, and the strength set is calculated according to the CARMA algorithm, and the strength set is combined together.

There are many attributes in the original four-level performance data, and attributes that are irrelevant, weakly relevant, or redundant with the research are deleted. In the end, the attributes we selected are the four-level total score (S), the total test score (T), the listening score (L), the reading score (R), and the writing score (W). According to the previous analysis of the student's related information factors, the related table is obtained. Then, the parameters are set; the fitness function method used is $f(x) = S(x) + B(x)$; the minimum support is 0.21, the minimum confidence is 0.81; the initial number of the population is set to 150; the mutation rate is 0.14; the crossover probability is 0.9.

4.2. Experimental Results. From Table 1, it can be seen that according to the combination of the genetic algorithm and the CARMA algorithm, the student performance mining found that there is a correlation between student absenteeism, usual performance, family status, and English performance.

The performance analysis of the improved algorithm and the CARMA algorithm is as follows. The number of excavations affects the efficiency of excavation. The memory is 512 Mb, the number of division levels is 6, and the algorithm is realized by MATLAB programming. The result is shown in Figure 7.

It can be seen from Figure 7 that the improved algorithm and the classic CARMA algorithm increase the mining time as the number of mining increases. The operating efficiency of the improved CRRMA algorithm has been improved, which is better than the improved algorithm in literature [24] and better than the classic CARMA algorithm, which shows that adding the genetic algorithm to an algorithm can improve the operating efficiency of the algorithm.

In order to investigate the accuracy of the improved algorithm in data mining, the existing data are obtained from the student information database. Mining data quality analysis is performed according to the formula as follows:

$$\eta = \frac{a}{a+b}\%. \quad (13)$$

In the above formula, η represents the mining success rate, a represents the number of successful mining, and b represents the number of failed mining. The excavation results are shown in Table 2.

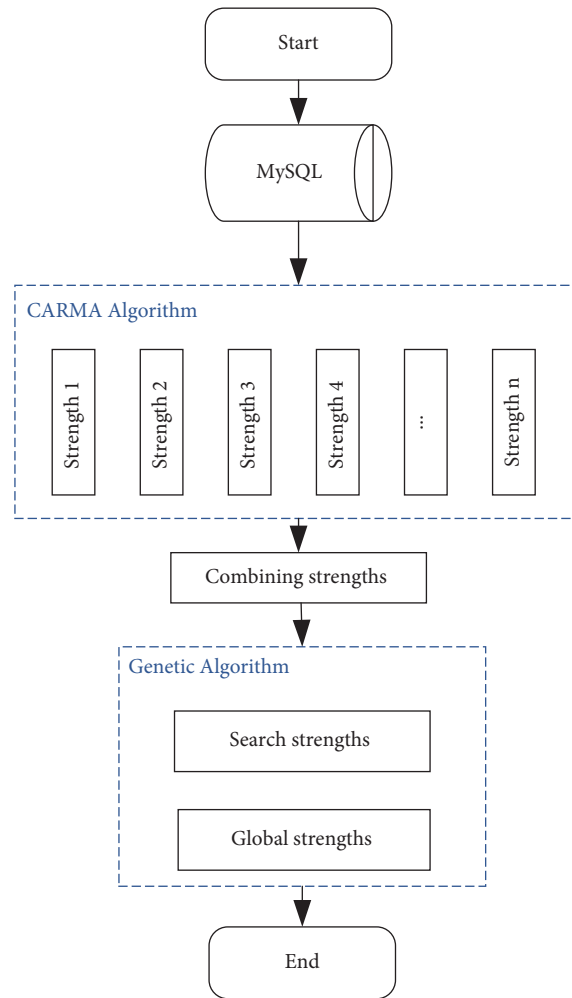


FIGURE 6: CARMA algorithm flowchart based on genetic algorithm.

Input: obtain the data set from the MySQL database and set the corresponding parameters of the genetic algorithm
 Output: output all frequent item set A

- (1) initial value assignment corresponding to genetic algorithm
- (2) for each I to T , do
 - //Find out frequent 1-items from the future data set and store the data in order
 - //Initial population
 End
- (3) while the population number in the CARMA algorithm is not 0
 - $A \leftarrow A \cup$ data set
 - For each two-unit (parent1, parent2) \in data set
 - Number of groups $N \leftarrow$ number of groups \cup cross (parent 1, parent 2)
 - End
 - For each unit \in number of groups
 - //Perform mutation operation to get the candidate set
 - $A \leftarrow A \cup$ data set
 - End
- (4) Output result

ALGORITHM 1: CARMA algorithm steps based on genetic algorithm.

TABLE 1: Part of student's English performance-related collection.

Number	Rule	Support	Confidence
1	The title of teacher and the quality of class (good grades)	0.342	0.859
2	The gender of teacher and the number of homework (average grades)	0.453	0.871
3	Teacher's length of service and degree of the course (good grades)	0.234	0.882
4	The quality of class and other subjects (average grades)	0.512	0.832
5	Student absent and Daily test (bad grades)	0.389	0.912
6	The number of homework and Family status (excellent grades)	0.276	0.903

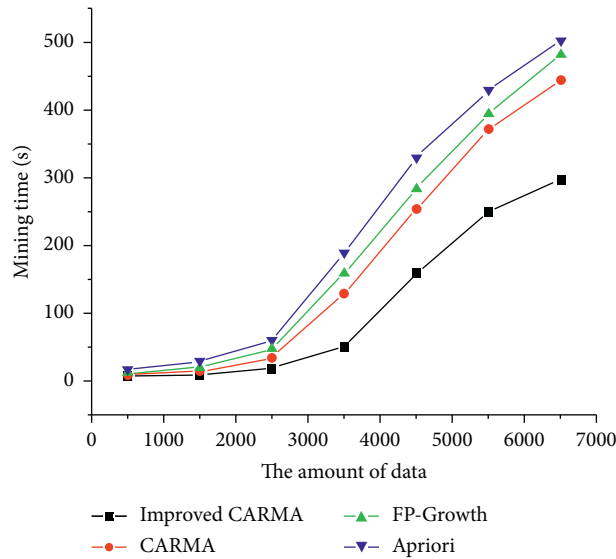


FIGURE 7: Impact of data volume on mining efficiency.

TABLE 2: Comparison of improved algorithm and CARMA algorithm mining quality.

	CARMA algorithm	Improved CARMA algorithm
Number of successes	10,006	10,023
Number of failures	844	202
Success rate	92.221	97.985
Failure rate	7.779	2.015

It can be seen from Table 2 that the mining accuracy rate of the improved algorithm reaches 97.985%, and the mining accuracy rate before the improvement is 92.221%, indicating that the improved algorithm can improve the accuracy of mining.

5. Conclusion

This paper uses data mining technology to analyze students' English scores. Aiming at the influence of many factors on the students' English scores, the analysis is realized by using the association rule algorithm. This paper is based on association rules. The students' English scores are analyzed and applied, and the main tasks are as follows:

- (1) At present, the problem of the CARMA algorithm is low operating efficiency. The combination of the genetic algorithm's crossover, mutation, and CARMA algorithm realizes the fast search of the

algorithm. The simulation results show that the operation performance of the algorithm is greatly improved after the crossover and mutation operations in the genetic algorithm are applied to the CARMA algorithm. The simulation results show that the mining accuracy of the improved algorithm is 97.985%, and the mining accuracy before the improvement is 92.221%, indicating that the improved algorithm can improve the accuracy of mining.

- (2) By comparing the mining time of the improved CARMA algorithm, the traditional CARMA algorithm, the FP-Growth algorithm, and the Apriori algorithm, the results show that when the number is 6,500, the mining efficiency of the improved CARMA algorithm is twice that of the other three algorithms. As the amount of data increases, the effect of improving mining efficiency gradually increases.

- (3) By using the improved CARMA algorithm to analyze students' English performance, it is found that the quality of student performance is strongly related to the quality of daily homework, and if it is related to the teacher's gender, professional title, etc., it is recommended that schools should pay more attention to homework during the teaching process.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] S. Natek and M. Zwilling, "Student data mining solution-knowledge management system related to higher education institutions," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6400–6407, 2014.
- [2] V. Kumar and C. Anupama, "An empirical study of the applications of data mining techniques in Higher education [J]," *International Journal of Advanced Computer ence& Applications*, vol. 2, no. 3, pp. 80–95, 2011.
- [3] D. Abdulazizalhamadi and M. Sabih Aksoy, "Data mining in education- an experimental study," *International Journal of Computer Applications*, vol. 62, no. 15, pp. 31–34, 2013.
- [4] K. Mahata and M. Fu, "On the indirect approaches for CARMA model identification," *Automatica*, vol. 43, no. 8, pp. 1457–1463, 2007.
- [5] R. Agrawal, T. Imieliski, and A. Swami, "Mining association rules between sets of items in large databases[C]//," in *Proceedings of the 1993 ACM SIGMOD International Conference on Management Data*, no. 2, pp. 207–216, ACM, New York, NY, USA, June 1993.
- [6] K. Korotkiewicz, M. Ludwig, and M. Stötzel, "State forecasting in smart distribution grids: a modular approach using CARMA algorithm[J]," in *Proceedings of the 24th International Conference on Electricity Distribution*, Glasgow, UK, 2017.
- [7] J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, "Association rule mining to detect factors which contribute to heart disease in males and females," *Expert Systems with Applications*, vol. 40, no. 4, pp. 1086–1093, 2013.
- [8] I. Frias-Blanco, J. Del Campo-Ávila, G. Ramos-Jimenez, and R. Morales Bueno, "Online and non-parametric drift detection methods based on Hoeffding's bounds[J]," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 3, pp. 810–823, 2014.
- [9] A. Bifet, J. Zhang, W. Fan et al., "Extremely fast decision tree mining for evolving data streams[C]//," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1733–1742, ACM, Halifax NS, Canada, 13 August 2017.
- [10] C. Manapragada, G. I. Webb, and M. Salehi, "Extremely fast decision tree[C]//," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1953–1962, ACM, London United Kingdom, 19 August 2018.
- [11] R. Pecori, P. Ducange, and F. Marcelloni, "Incremental learning of fuzzy decision trees for streaming data classification[C]//," in *Proceedings of the 11th Conference of the European Society for Fuzzy Logic and Technology*, no. s.n., pp. 748–755, Atlantis, 2019.
- [12] S. Shalev-Shwartz, Y. Singer, and N. Srebro, "Pegaso: rimal estimated sub-gradient solver for SVM[J]," *Mathematical Programming*, vol. 127, no. 1, pp. 3–30, 2011.
- [13] T. LE, T. D. Nguyen, V. Nguyen, and D. Phung, "Approximation vector machines for large-scale online learning[J]," *Journal of Machine Learning Research*, vol. 18, no. 1, pp. 3962–4016, 2017.
- [14] A. Kobren, N. Monath, and A. Krishnamurthy, "A hierarchical algorithm for extreme clustering[C]//," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 255–264, ACM, Halifax NS Canada, August 2017.
- [15] D. Mautz, W. Ye, C. Plant, and C. Bohm, "Discovering non-redundant K-means clusterings in optimal subspaces[C]//," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1973–1982, ACM, Germany, 19 August 2018.
- [16] K. Djouzi and K. Beghdad-Bey, "A review of clustering algorithms for big data[C]//," in *Proceedings of the 2019 International Conference on Networking and Advanced Systems (ICNAS)*, no. S.I., pp. 1–6, IEEE, June 2019.
- [17] A. Hadian and S. Shahrivari, "High performance parallel K-means clustering for disk-resident dataset on multi-core CPUs[J]," *The Journal of Supercomputing*, vol. 69, no. 2, pp. 845–863, 2014.
- [18] D. Melo, S. Toledo, F. Mourão et al., "Hierarchical density-based clustering based on GPU accelerated data indexing strategy[J]," *Procedia Computer Science*, vol. 80, pp. 951–961, 2016.
- [19] N. Monath, K. A. Dubey, G. Guruganesh et al., "Scalable hierarchical agglomerative clustering[C]//," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 1245–1255, ACM, Singapore, 14 August 2021.
- [20] S. Gole and B. Tidke, "ClustBIGFIM-frequent itemset mining of big data using pre-processing based on mapreduce framework[J]," *International Journal in Foundations of Computer Science & Technology (IJFC-ST)*, vol. 5, no. 3, pp. 79–89, 2015.
- [21] S. Rathee, M. Kaul, and A. Kashyap, "R-A priori: An efficient apriori based algorithm on spark[C]//," in *Proceedings of the 8th workshop on Ph d Workshop In information and Knowledge Management*, pp. 27–34, ACM, Melbourne, Australia, 19 October 2015.
- [22] S. Rathee and A. Kashyap, "Adaptive-Miner: An efficient distributed association rule mining algorithm on Spark[J]," *Journal of Big*, vol. 5, no. 1, pp. 1–17, 2018.
- [23] J. Heaton, "Comparing dataset characteristics that favor the Apriori, Eclat or FP-Growth frequent itemset mining algorithms[C]//," in *Proceedings of the Southeast Con2016*, no. S.I., pp. 1–7, IEEE, Norfolk, Virginia, USA, 3 April 2016.
- [24] Y. Djenouri, D. Djenouri, A. Belhadi, and A. Cano, "Exploiting GPU and cluster parallelism in single scan frequent itemset mining [J]," *Information Sciences*, vol. 496, pp. 363–377, 2019.

Research Article

Application of Image Mosaic Technology in Tai Chi Animation Creation

Yajun Pang 

College of Physical Education, Luoyang Institute of Science and Technology, Luoyang, Henan 471023, China

Correspondence should be addressed to Yajun Pang; yajun.pang@lit.edu.cn

Received 7 December 2021; Revised 28 December 2021; Accepted 30 December 2021; Published 13 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yajun Pang. 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.

Panorama can reflect the image seen at any angle of view at a certain point of view. How to improve the quality of panorama stitching and use it as a data foundation in the “smart tourism” system has become a research hotspot in recent years. Image stitching means to use the overlapping area between the images to be stitched for registration and fusion to generate a new image with a wider viewing angle. This article takes the production of “Tai Chi” animation as an example to apply image stitching technology to the production of realistic 3D model textures to simplify the production of animation textures. A handheld camera is used to collect images in a certain overlapping area. After cylindrical projection, the Harris algorithm based on scale space is adopted to detect image feature points, the two-way normalized cross-correlation algorithm matches the feature points, and the algorithm to extract the threshold T iteratively removes mismatches. The transformation parameter model is quickly estimated through the improved RANSAC algorithm, and the spliced image is projected and transformed. The Szeliski grayscale fusion method directly calculates the grayscale average of the matching points to fuse the image, and finally, the best stitching method is used to eliminate the ghosting at the image mosaic. Data experiments based on Matlab show that the proposed image splicing technology has the advantages of high efficiency and clear spliced images and a more satisfactory panoramic image visual effect can be achieved.

1. Introduction

Tai Chi is one of China’s precious cultural resources and one of the important cultural industries to promote regional economic development. This article uses animation to promote this cultural product. When using 3D animation production software to design anime character models, in order to save manpower and time, low-precision models can be used. However, in order to achieve exquisite and realistic visual effects for the characters, the real character images are used as the model textures when making the model textures. To make up for the shortcomings of low-precision models. However, the collected real person images need to be spliced into a two-dimensional plane 360° panoramic view according to the requirements of the model UV distribution map, which can be completed by using image splicing technology. Image stitching technology is used in many areas of life. For example, in the field of medical research,

with the increasing use of microscopes, pictures under high magnification microscopes are likely to separate the areas that doctors want to observe. The image mosaic technology can be adopted to generate a new mosaic to help doctors solve the problem of not being able to pass one. Other application fields include military field, aerospace field, mechanical design field, and so on. This article will take a group of real person head images as an example to describe the process of using image stitching technology to make 3D model textures [1–5].

Image mosaic technology is an important research topic in the fields of virtual reality, computer graphics, and image processing. In the field of animation, it is used to make animated model textures to save costs and improve the speed and accuracy of making model textures. In the field of TIP (tour into the picture), panoramic images are generated to provide an important basis for generating dynamic three-dimensional images in two-dimensional still images.

Image stitching technology (see Figure 1) is the registration and fusion of several adjacent images or photos with overlapping areas to form a 360-degree or wide-view panoramic image. The image stitching algorithm mainly focuses on the method based on the area correlation and the feature. The method based on the area correlation determines the variance function through the similarity of the images and determines the transformation relationship between the images through the function. This method requires the transformation amplitude between the images to be small, and the method is easy to be affected by care conditions and has a large amount of computer. The feature-based method is to extract the image features and determine the corresponding relationship between the features to find the transformation relationship between the images. This method is relatively stable, fast, and has a wide range of adaptation [6–8].

As one of the important contents of image stitching, image registration is mainly divided into the following.

1.1. Registration Based on Image Grayscale. This type of method directly uses image gray information to construct a similarity function for similarity calculation. Commonly used methods are the projection matching method proposed by Fuh CS and the mutual information matching method proposed by P. Viola. The principle of the image registration algorithm based on gray information is simple, but the amount of calculation is large, and because it calculates the gray information of pixels, it is very sensitive to noise and light.

1.2. Image Registration Based on Transform Domain. Transform the picture into the frequency domain, and then perform the corresponding transformation processing to realize the image registration, such as the phase correlation matching method based on Fourier transform proposed by Srinivasa and Reddy. The time delay in the time domain corresponds to the phase change in the frequency domain, and the corresponding change in the amplitude is relatively simple, so it is generally used in affine conversion image registration, but it is not suitable for complex graphics conversion forms.

1.3. Image Registration Based on Features. Image features include feature points, contour features, regional special diagnosis structure, edge features, and so on. In addition, there are many commercial image stitching software, such as open-source software Hugin, Apple's QuickTime VR system, Microsoft's Digital Image, and so on.

The domestic image splicing technology started later than Europe and America, but it has developed rapidly in recent years. In the field of edge extraction, Wen Ting et al. comprehensively considered the effects of image gradient features, phase features, and noise and used directional energy and the brightness gradient calculated by the histogram difference operator as image features for edge detection; Sun Shuyi et al. used wavelet transform extracts

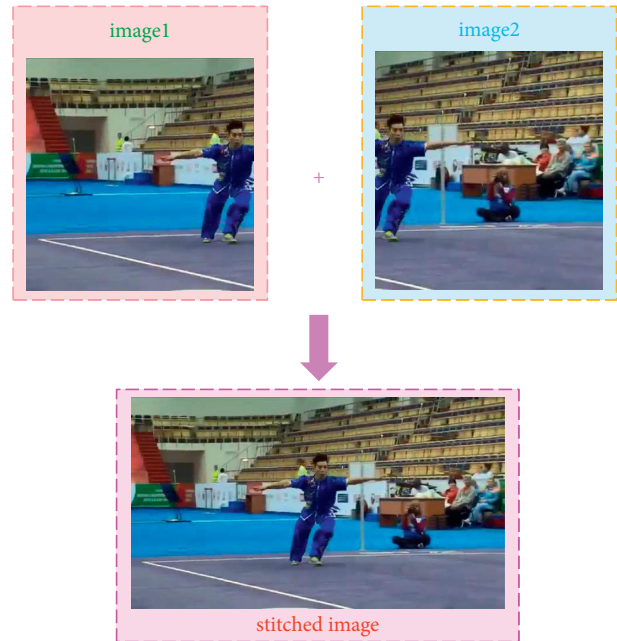


FIGURE 1: Image stitching technology.

edge features, and then uses manual selection for registration. In terms of feature point extraction, the feature point extraction based on the variational B-spline filter combined with the Harris algorithm proposed by Zhang Yong et al. has significantly improved the time complexity of the algorithm; the tracking grouped Harris proposed by Xu Xianfeng et al. Laplace algorithm feature point extraction has the advantage of reducing feature redundancy; Qiu Guoqing and others improved the Harris algorithm to avoid the k value in the corner response function and the selection of the threshold T during the detection process [9–12].

Although many scholars at home and abroad have done a lot of research on feature extraction and matching, and have achieved obvious results, the research work on how to improve the speed of feature extraction, improve matching accuracy, reduce redundancy, and enhance robustness is still ongoing. The author thinks that it is difficult to find an algorithm that is suitable for all image stitching, after all, the fields of image stitching are not the same. In the medical field, the emphasis is on the quality of spliced images, because if mismatches, missing matches, and a little ghosting occur, they will have a greater impact on the doctor's judgment; in the aerospace field, the emphasis is on the robustness of spliced images. Great because the objects in the celestial pictures taken are moving and the light is different. In the smart tourism system, due to the high quality of pictures, the focus of algorithm improvement should be to reduce redundancy and reduce the time complexity of the stitching algorithm [13–15].

The main steps of panoramic image production are based on feature points. First, the image with a certain overlapping area is collected, and then the image is pre-processed. According to the stitching method in this article, the image is cylindrically projected, then extract feature

points, match feature points, image transformation, and finally generate a panorama through image fusion. Image acquisition is divided into three methods: the camera is placed on a pulley to shoot in parallel, the camera is placed on a tripod to take a fixed rotation, and the camera is held in hand to shoot. This article uses the most common and flexible handheld camera shooting method. This method is widely used and has practical research significance. It is difficult to stitch together images collected by a handheld camera because the camera movement is complicated during the shooting process, and there is a horizontal and vertical offset. There may be a certain exposure difference between adjacent images and a small range of object movement, or even a certain amount in the rotation of the angle and the scaling of the small scale (the scale is the size of the object or feature) [16–23]. The various objects or features of the collected images need to be studied within a specific scale to make sense. A panoramic image is a representation of a discrete image processed into a continuous image. In order not to destroy the visual consistency between various objects in the image and maintain the spatial constraint relationship in the actual scene, the collected discrete images need to be projected onto standard coordinates. According to the form of its projection surface, the panorama can be divided into three types: spherical surface, cube surface, and cylindrical surface. Because the cylindrical projection image is easy to obtain, the image quality is uniform, and the detail is high, it is widely used. This article is an image stitching algorithm that uses a cylindrical surface as the projection surface. After the feature points are extracted, the two-way normalized cross-correlation method is used for feature point matching. The transformation matrix calculation is to calculate the spatial geometric transformation of two images to be spliced according to the feature point matching point set. The transformation matrix model used in this article is an eight-parameter projection transformation model. The model describes the translation, rotation, scaling and regularity of the image stretching also includes irregular stretching, that is, parallel lines cannot remain parallel after transformation. After the image is cylindrically projected, an image matching operation is required, that is, the images of adjacent overlapping areas are spliced and merged. The splicing method based on feature points extracts image feature locations by describing features. This paper adopts a Harris corner detection feature point algorithm with scale spatial information. After the image projection transformation, the image needs to be fused, that is, the coordinates of the reference image and the target image are merged into one coordinate. The directly stitched image often has problems such as grayscale difference and ghosting, which requires the stitched image perform grayscale adjustments and eliminate ghosting operations [24–29].

2. Image Stitching

When the “Tai Chi” character image is collected, the postimage stitching and the UV distribution of the character model should be fully considered. In order to obtain a 360° panoramic image of the “Tai Chi” character’s head, when the

image is collected, the “Tai Chi” character’s head is taken as the central axis, and the camera is rotated around the central axis to take a picture. During the shooting process, continuous shooting from multiple angles is required. The lens must be aimed at the head of the person, and there must be a certain degree of overlap between the consecutive images taken. The way of image acquisition is shown in Figure 2.

Image stitching is a complex process, and the main steps of the process can be summarized as: image preprocessing, image projection transformation, image registration, and image fusion.

2.1. Image Preprocessing. For cylindrical panoramas, we require the camera to remain still on the horizontal line when capturing images, try to avoid the pitch and tilt of the lens, and perform 360-rotation shooting in the horizontal direction. If conditions permit, place the camera on a fixed tripod, and the effect will be better. The number of pictures taken determines the workload of stitching. The more pictures, the greater the workload. However, if the visual range of the width of each picture is expanded in order to reduce the stitching time, the overlap area between adjacent pictures will be reduced, resulting in inaccurate feature matching and even stitching failure. Weighing the two sides, we generally require that the overlapping area between adjacent pictures be about 50%. Finally, the content to be shot is required to be immobile, that is, there are no moving people or objects; otherwise, it will cause distortion of the stitching graphics or even stitching failure.

Due to the changeable environment, the image will inevitably have noise points. These noise points may cause interference during feature extraction and affect the matching effect and efficiency. Therefore, before the projection transformation, the image is denoised first. Properties such as brightness, color, shape, and grayscale of a set of continuously acquired images vary due to human or natural factors. Therefore, it is necessary to preprocess the problems of low contrast, deformation, geometric distortion, and so on that appear in the image. This will not only improve the display quality of the image but also it can ensure the smooth progress of the subsequent image registration process.

2.2. Image Projection. Why do we need to project? Can the overlapping captured images be directly stitched? It can be spliced, but the effect is not ideal. Because the image is the projection of the real scene in two-dimensional coordinates, directly splicing the image cannot meet the human’s requirement for visual consistency. Simply put, the pictures taken are horizontal plane imaging, and the pictures we need are cylindrical imaging. Direct stitching will produce distortion.

The biggest feature of the cylindrical panorama is that it has no top and no bottom and can only be used for small-angle pitch operations. This article chooses cylindrical panorama as the research content for the following reasons:

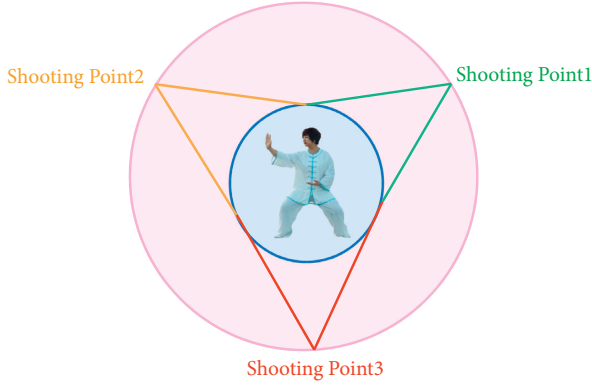


FIGURE 2: The way of image acquisition.

- (1) The picture collection of the cylindrical panorama is simple
- (2) The edge distortion of the cylindrical panorama is small, and the pixels are evenly distributed
- (3) The coordinate transformation of the cylindrical panorama is relatively simple, and there are no edges and fixed points in the cube panorama, so the determination of coordinates and the difficulty of stitching are reduced

Cylindrical panoramas need to project the plane image onto a cylindrical body with the camera focal length f as the radius. At the same time, the focal length of the camera needs to be used in the plane and cylindrical coordinate transformation and the calibration of the camera coordinates. The camera focal length f is a panorama, which is an important parameter in the generation process. But the focal length parameters of most cameras are not given, especially for pictures taken with mobile phones on the road. Therefore, before the cylindrical projection of the image, we must first estimate the focal length of the shooting camera used.

The images collected according to the method shown in Figure 2 were taken at different angles. In order to maintain the spatial constraint relationship of the actual objects, the collected images need to be projected on a standard coordinate system. Otherwise, if you directly control these images and the splicing is performed, the visual consistency between various objects in the actual scene will be destroyed, and people's visual requirements will not be met. Because the shape of the head profile used in this article is like a cylinder, the projected image has nothing to do with its projected position on the surface of the cylinder. A cylindrical projection model as the standard coordinate space is selected in this article.

As shown in Figure 3, the coordinate system is established with O as the origin, I is an original image taken by the camera, where point $P(x, y)$ is any pixel on the original image I , then the pixel point P is the coordinates in the camera coordinate system xyz can be expressed as

$$\left(x - \frac{W}{2}, y - \frac{H}{2}, -f\right). \quad (1)$$

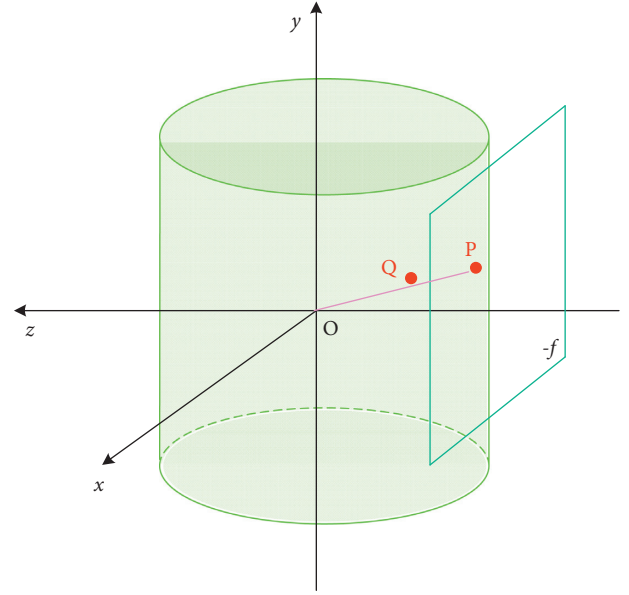


FIGURE 3: Coordinate system.

Here, W is the width of the original image I , and H is the height of the original image I . If the center of the cylinder is used as the origin of the camera's coordinate system, and the camera's pixel focal length f is used as the radius of the cylinder, then in a cylindrical panoramic image, the projection point $Q(x', y')$ can be expressed as

$$\begin{cases} x' = f \arctan\left(\frac{x - W/2}{f}\right) + f \arctan\left(\frac{W}{2f}\right), \\ y' = \frac{f(y - H/2)}{\sqrt{(x - W/2)^2 + f^2}} + \frac{H}{2}. \end{cases} \quad (2)$$

Here, f is the focal length and satisfies the following formula:

$$f = \frac{W}{2 \tan(hf_{ov}/2)}. \quad (3)$$

It can be concluded from the formula (2) that the cylindrical projection algorithm can prevent the object from being deformed in the vertical direction.

2.3. Image Registration. Image registration uses similarity measures to calculate spatial transformation parameters and transforms two or more pictures of the same scene with a certain degree of similarity from different perspectives and at different times to the same coordinate system to obtain a spliced picture. Image registration has always been the focus of image mosaic technology research, and scholars at home and abroad have also proposed a variety of image registration algorithms. In general, image registration must include the following basic content:

- (1) Feature space, that is, the collection of features of the image to be matched, such as grayscale features, statistical features (moment invariants, center), edges, corners, etc.: when selecting features, attention should be paid to the number of selected features, too much will increase the amount of matching, and too little will cause matching failures.
- (2) Selection of similarity measures: it is used to measure the similarity between matching features. It is the most critical step of registration and directly determines the accuracy of matching.
- (3) Search space, that is, all possible transformation spaces.
- (4) The best space, which is to find the transformation parameters with the highest similarity in the search space.

Due to the different shooting time of the two images during the stitching process, the gray value of the same scene is not necessarily the same, so stitching will occur at the stitching. The purpose of image fusion is to eliminate seams, make the image naturally excessive, and maintain the good spectral characteristics of the image itself (Figure 4).

The basic principle of image registration is to splice multiple pictures containing the same image area and use a certain matching method for the same image area to determine the splicing position between two adjacent images. Based on the characteristics of a person's head image with rich edge information, texture information, and feature point information, this article adopts the Harris corner detection algorithm based on feature point matching. The corner detection formula is

$$E(\Delta x, \Delta y) = \sum_{x,y} w(x, y) [I(x + \Delta x, y + \Delta y) - I(x, y)]^2, \quad (4)$$

where $w(x, y)$ is the window function and $I(x + \Delta x, y + \Delta y) - I(x, y)$ is the gradient value of the image grayscale. For each small displacement $(\Delta x, \Delta y)$, the bilinear approximation can be expressed as

$$E(\Delta x, \Delta y) \cong [\Delta x, \Delta y] M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}. \quad (5)$$

Here, M can be expressed by the following formula:

$$M = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}. \quad (6)$$

Let λ_1 and λ_2 be the two eigenvalues of the matrix M , representing the curvature of the local autocorrelation function. E can be approximated as a local cross-correlation function, describing the shape at this point. When the values of both λ_1 and λ_2 are small, if the window moves in any direction, the change in the value of E is small; when one of λ_1 and λ_2 is large, and the other value is small, if the window moves along the edge direction, the change in E value is also small, and if the window moves in the direction perpendicular to the edge, the change in E value is greater; and

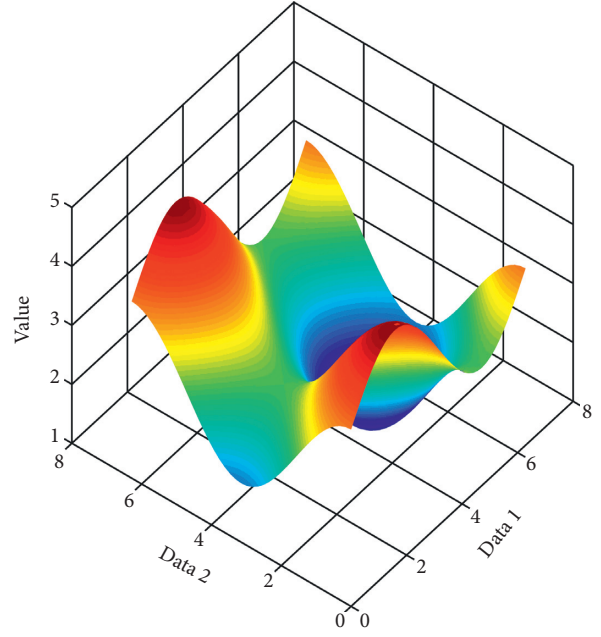


FIGURE 4: Value with different data.

when both λ_1 and λ_2 are both large, the moving in any direction will cause the E value to increase sharply. The value variation is plotted in Figure 5. According to the above situation, the response function used to calculate the corner point in actual application can be written as

$$R = \det(M) - k \text{Trace}^2(M), \quad (7)$$

where

$$\det(M) = \lambda_1 \lambda_2, \text{Trace}(M) = \lambda_1 + \lambda_2. \quad (8)$$

When the determinant of the matrix M is large, it indicates that it is an edge or a corner; when the sum of the main diagonals of the matrix M in a certain area is large, it indicates that it is an edge. k generally takes an empirical value of 0.04.

The Harris corner detection algorithm can extract the feature point information used to register the two images. However, the extracted feature point information cannot be used for matching. It is also necessary to measure the similarity of these feature point information. The performance metric determines the relevant characteristics of all registration tests. First, extract a correlation window of $(2N + 1) \times (2N + 1)$ size in the two images with each feature point as the center, and then use each feature point in the reference image as a reference point, and proceed according to the feature point in the current image. Search sequentially and perform block matching. The search area can be specified during matching, and its range can also be specified according to experimental experience. With this method, the search area of the corresponding feature point can be changed from the entire image to a window of a specified size, which can greatly reduce the calculation workload. This article uses the normalized cross-correlation (NCC) algorithm to calculate the correlation coefficient between

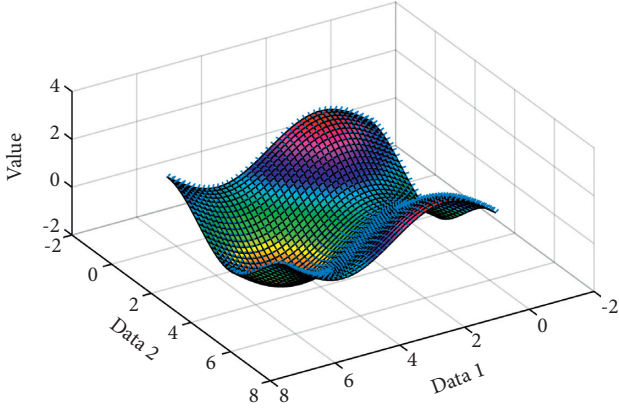


FIGURE 5: Value variation.

two correlation windows. After the rough matching of feature points above, there are still a small number of false feature pairs. If the least squares method is used to estimate the model parameters directly, it will bring a relatively large registration error. Therefore, this paper uses the RANSAC method to achieve fine matching. This algorithm can effectively exclude outliers in feature matching.

2.4. Image Fusion. After the image is spliced, due to the different perspectives, resolutions, etc., of the original image plus the influence of external factors such as lighting, the spliced image will produce blur, noise, or ghost in the overlapped part, and the splicing will occur at the same time. There may also be seams at the borders. Therefore, it is necessary to perform fusion processing on the spliced images. In this article, the multiresolution fusion method is adopted, and the implementation process is as follows:

- (1) Create a pyramidal hierarchical structure of the original image to obtain the low-pass layer of each image:

$$G_l(x, y) = \sum_{m,n=-2}^2 w(m, n)G_{l-1}(2x + m, 2y + n). \quad (9)$$

Here, G_0 is the original image.

- (2) Through the obtained low-pass layer of each image, the bandpass layer of the corresponding image can be decomposed:

$$L_l(x, y) = G_l(x, y) - 4 \sum_{m,n=-2}^2 G_l\left(\frac{2x+m}{2}, \frac{2y+n}{2}\right). \quad (10)$$

- (3) Perform the image fusion operation in the bandpass layer of each image separately and use the weighted average method to achieve, then for the current L_k layer, there are

$$L_{kout}(x, y) = \frac{\sum_{i=0}^{i=N-1} L_{kli}(x', y')w_i(x')w_i(y')}{\sum_{i=0}^{i=N-1} w_i(x')w_i(y')}. \quad (11)$$

After this step, the bandpass space corresponding to the output image is obtained.

- (4) Combine the obtained bandpass layers to obtain the final stitched image:

$$G_{out} = \sum_{k=0}^N L_{out}. \quad (12)$$

It can be seen through experiments that the fused image is clear, smooth, and seamless.

3. Image Feature Extraction and Matching

The feature of an image is the abstract expression of the pixels or collections of pixels in the image, and it is the most basic attribute that distinguishes the image. It can describe image information with fewer pixels. The feature extraction of the image is chosen because it can avoid the influence of noise, grayscale, and other external factors to a greater extent.

The image features include feature lines, feature faces, feature points, and so on. Special line: the feature line mostly refers to the edge features of the image, including contours and arcs. Usually, edge detection algorithms are used to detect the characteristic curve. The error variation is shown in Figure 6:

Feature points: feature points are points that reflect the types of features or geographic distribution characteristics of the area. Generally speaking, feature points refer to corner points (currently, there is no clear mathematical definition) or extreme points in the calculated values of other feature extraction functions. The corner points where the boundary direction changes significantly have the greatest probability.

Feature surface: it refers to the area block that contains the significant information in the image, also called the feature domain. Most of the feature surface extraction is the obvious closed area, and the commonly used extraction method is the segmentation method. The biggest difficulty of image stitching technology based on feature surface lies in the extraction of feature regions.

The matching feature point pairs obtained by the NCC algorithm will have a lot of errors. One is caused by the inaccurate positioning of the feature points. This effect can be reduced when using the least squares method to solve the transformation matrix, and the other is due to interference such as noise. If there are many mismatches, it will have a great influence on the matrix of graph transformation. Therefore, after the NCC matching is completed, the matching pair will be “purified.” Here we introduce the RANSAC algorithm. The image process is shown in Figures 7 and 8.

The RANSAC algorithm is a mathematical model that uses an iterative algorithm to estimate parameters. It is an algorithm with a high fault tolerance rate, combined with the NCC algorithm, it can effectively bring out the

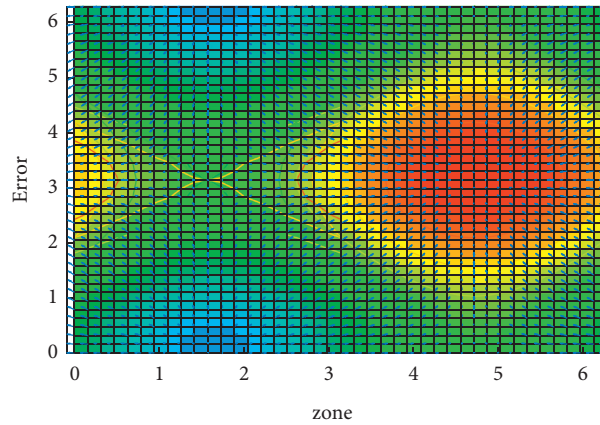


FIGURE 6: Error variation.

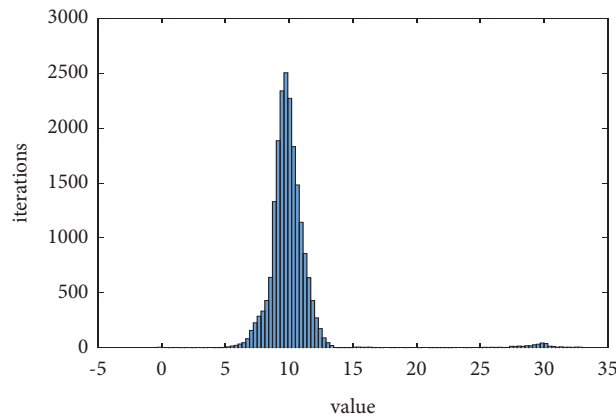


FIGURE 7: Image process.

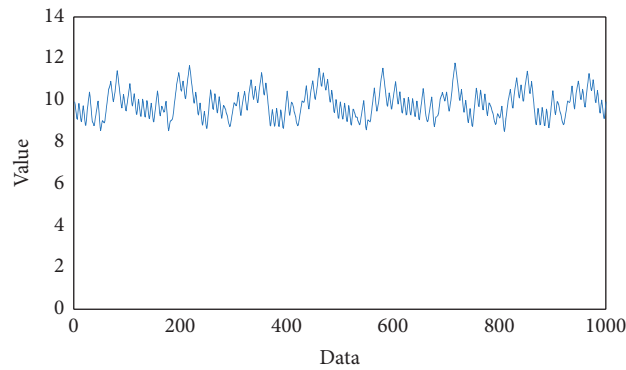


FIGURE 8: Process data.

external points. The idea of the RANSAC algorithm is to find the transformation model that contains the most interior points through iteration. Taking straight line fitting as an example, we need to find a fitted straight line that contains as many points in the figure as possible. We first select two points, find a straight line, and then take a certain threshold t , mark the points with a distance less than t from the straight line as interior points, and then

find a straight line based on the interior point set obtained this time, and use t to judge the same. How many points belong to the interior points of the straight line at this time, until a certain sampling makes the number of interior points the largest, and the straight line at this time is the desired one.

For images with a large amount of data, we continue to sample and find the transformation matrix. The H matrix



FIGURE 9: The stitched image.

determined by every 4 matching points needs to be verified one by one, which is a lot of work. In this article, when using the RANSAC algorithm to purify and calculate the H matrix, an improvement is made: 4+2 pairs are selected for the initial matching point, and the H matrix determined by the 4 feature points is determined by the other 2 pairs of feature points. If they are not all interior points, then just discard it and choose 4+2 pairs of matching points.

The images used in the experiment are images collected with a handheld mobile phone under the conditions of fixed-point rotation and translation under outdoor lighting, and the overlap rate of the collected images is about 20%. After the image is collected, the image is first cylindrically projected, the image feature points are extracted using the Harris algorithm with invariant scale space, and the two-way normalized cross-correlation method is used for matching, the feature points are extracted iteratively to remove mismatches, and the projection model is estimated using the RANSAC algorithm, and finally adopt the Szeliski grayscale fusion method. The method of direct gray-level averaging of matching points and the best stitching line completes the mosaic processing of the panorama. Figure 5 is the final stitching effect picture, and the stitching image quality can meet the requirements of TIP background image and animation model texture. The stitched image is shown in Figure 9.

4. Conclusion

Image stitching technology is a popular research direction in the field of digital image processing, and it has attracted many scholars to study it in recent decades. Among them, image registration is its key technology. For image registration technology, people generally hope to improve from three aspects, one is the accuracy of feature extraction, the other is the extraction speed, and the third is the accuracy of matching.

In the creation of Tai Chi animation, the pictures that need to be spliced are often portrait pictures, and to achieve a good publicity effect, the quality of the pictures must be relatively high, and there may be more features, which increases the workload of feature matching. Starting from the general direction of saving time, the article sacrifices a small amount of time to reduce the accumulation of feature points that are easy to appear in the classic Harris algorithm, reduces redundancy, reduces the workload for feature matching of the following pictures, and proposes to increase two pairs of matching points. The prejudgment matching matrix greatly reduces the calculation time of the RANSAC

algorithm, especially the transformation model calculation for high-quality tourist scenery pictures.

In my country, the National Tourism Administration has included “smart tourism” in the “Twelfth Five-Year Tourism Development Plan.” In a policy-supported environment, the rapid development of Internet of Things technology, cloud computing, modern communication technology, and virtual reality technology related to smart tourism has also provided technical support for it. The popularity of smart phones and tablet computers has also provided them with strong hardware support and user base. Therefore, the development prospects of smart tourism are very good.

The formation of a cylindrical panorama is simpler than that of a spherical panorama, but it cannot be transformed from the top view and the top view, and when roaming, the distortion on both sides is obvious. The matching strategy in this paper is pairwise registration. For multiimage splicing, the matching process is repeated many times, which wastes a lot of time. How to realize the registration and fusion of multiple images together is also a problem that the author considers. The last is the loading of recommended information. Now in the system, the panorama can only be browsed, without any information displayed. Because the cruise ship is very large, the author considers the actual consumption information of previous passengers and the click-through rate of each panorama as the criterion and adds recommended stars to the panorama to make the purpose of tourist’s cruise clearer.

Data Availability

The data set can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Science and Technology Projects of Henan Science and Technology Department under projects “Assessment of Action in Sports Video Based on Multifeature Fusion Number” (182102310041) and “Assessment of Taichi Action Based on Vision Transformer” and Sanmenxia Vocational And Technical College Project under project “Assessment of Action in Low-Quality Large-Displacement Sports Videos Number” (SZYGCCRC-2020-005).

References

- [1] Y. Yu, C. Yang, Q. Deng, T. Nyima, S. Liang, and C. Zhou, "Memristive network-based genetic algorithm and its application to image edge detection," *Journal of Systems Engineering and Electronics*, vol. 32, no. 5, pp. 1–9, 2021.
- [2] Y. Ishida and S. Hashimoto, "Asymmetric characterization of diversity in symmetric stable marriage problems: an example of agent evacuation," *Procedia Computer Science*, vol. 60, no. 1, pp. 1472–1481, 2015.
- [3] P. Zoha and R. Kaushik, "Image edge detection based on swarm intelligence using memristive networks," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1774–1787, 2018.
- [4] J. Pais, "Random matching in the college admissions problem," *Economic Theory*, vol. 35, no. 1, pp. 99–116, 2018.
- [5] J. J. Jung and G. S. Jo, "Brokerage between buyer and seller agents using constraint satisfaction problem models," *Decision Support Systems*, vol. 28, no. 4, pp. 291–384, 2020.
- [6] Y. Liu and K. W. Li, "A two-sided matching decision method for supply and demand of technological knowledge," *Journal of Knowledge Management*, vol. 21, no. 3, 2017.
- [7] J. Byun and S. Jang, "Effective destination advertising: matching effect between advertising language and destination type," *Tourism Management*, vol. 50, no. 10, pp. 31–40, 2015.
- [8] A. N. Nagamani, S. N. Anuktha, N. Nanditha, and V. K. Agrawal, "A genetic algorithm-based heuristic method for test set generation in reversible circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 2, pp. 324–336, 2018.
- [9] C. Koch and S. P. Penczynski, "The winner's curse: conditional reasoning and belief formation," *Journal of Economic Theory*, vol. 174, pp. 57–102, 2018.
- [10] C. K. Karl, "Investigating the winner's curse based on decision making in an auction environment," *Simulation & Gaming*, vol. 47, no. 3, pp. 324–345, 2016.
- [11] D. Ettinger and F. Michelucci, "Creating a winner's curse via jump bids," *Review of Economic Design*, vol. 20, no. 3, pp. 173–186, 2016.
- [12] J. A. Brander and E. J. Egan, "The winner's curse in acquisitions of privately-held firms," *The Quarterly Review of Economics and Finance*, vol. 65, pp. 249–262, 2017.
- [13] Z. Palmowski, "A note on var for the winner's curse," *Economics/Ekonomia*, vol. 15, no. 3, pp. 124–134, 2017.
- [14] B. R. Routledge and S. E. Zin, "Model uncertainty and liquidity," *Review of Economic Dynamics*, vol. 12, no. 4, pp. 543–566, 2009.
- [15] D. Easley and M. O'Hara, "Ambiguity and nonparticipation: the role of regulation," *Journal of European Economy*, vol. 22, no. 5, pp. 1817–1843, 2019.
- [16] P. Klibano, M. Marinacci, and S. Mukerji, "A smooth model of decision making under ambiguity," *Econometrica*, vol. 73, no. 6, pp. 1849–1892, 2005.
- [17] Y. Halevy, "Ellsberg revisited: an experimental study," *Econometrica*, vol. 75, no. 2, pp. 503–536, 2017.
- [18] D. Ahn, S. Choi, D. Gale, and S. Kariv, "Estimating ambiguity aversion in a portfolio choice experiment," *Working paper*, vol. 5, no. 2, pp. 195–223, 2019.
- [19] T. Hayashi and R. Wada, "Choice with imprecise information: an experimental approach," *Theory and Decision*, vol. 69, no. 3, pp. 355–373, 2010.
- [20] K. Zima, E. Plebankiewicz, and D. Wiczorek, "A SWOT Analysis of the Use of BIM Technology in the Polish Construction Industry," *Buildings*, vol. 10, no. 1, 2020.
- [21] P. Sun, B. Liu, and T. Sun, "Injury status and strategies of female 7-a-side rugby players in Anhui Province," *Sports Boutique*, vol. 38, no. 3, pp. 72–74, 2019.
- [22] P. Guild, M. R. Lininger, and M. Warren, "The association between the single leg hop test and lower-extremity injuries in female athletes: a critically appraised topic," *Journal of Sport Rehabilitation*, vol. 30, no. 2, pp. 320–326, 2020.
- [23] U. G. Inyang, E. E. Akpan, and O. C. Akinyokun, "A hybrid machine learning approach for flood risk assessment and classification," *International Journal of Computational Intelligence and Applications*, vol. 19, no. 2, Article ID 2050012, 2020.
- [24] Q. Liu, S. Du, B. Wyk, and Y. Sun, "Double-layer-clustering differential evolution multimodal optimization by speciation and self-adaptive strategies," *Information Sciences*, vol. 545, no. 1, pp. 465–486, 2021.
- [25] H. R. Medeiros, F. D. Oliveira, H. F. Bassani, and A. Araujo, "Dynamic topology and relevance learning SOM-based algorithm for image clustering tasks," *Computer Vision and Image Understanding*, vol. 179, no. FEB, pp. 19–30, 2019.
- [26] Y. Deng, D. Huang, S. Du, G. Li, and J. Lv, "A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis," *Computers in Industry*, vol. 127, Article ID 103399, 2021.
- [27] J. J. Chan, K. K. Chen, S. Sarker, J. Z. Hasija, and E. Vulcano, "Epidemiology of Achilles tendon injuries in collegiate level athletes in the United States," *International Orthopaedics*, vol. 44, no. 3, pp. 585–594, 2020.
- [28] W. Li, G. G. Wang, and A. H. Gandomi, "A survey of learning-based intelligent optimization algorithms," *Archives of Computational Methods in Engineering*, vol. 28, pp. 3781–3799, 2021.
- [29] G. G. Wang, A. H. Gandomi, A. H. Alavi, and Gong, "A comprehensive review of krill herd algorithm: variants, hybrids and applications," *Artificial Intelligence Review*, vol. 51, no. 1, pp. 119–148, 2019.

Research Article

Research on Teaching Practice of Blended Higher Education Based on Deep Learning Route

Yang Li,¹ Lijing Zhang,¹ Yuan Tian,¹ and Wanqiang Qi ²

¹Aviation University of Air Force, Changchun, Jilin 130022, China

²School of Automotive Engineering, Jilin Teachers Institute of Engineering and Technology, Changchun, Jilin 130022, China

Correspondence should be addressed to Wanqiang Qi; qwqwall@jlenu.edu.cn

Received 29 November 2021; Revised 15 December 2021; Accepted 24 December 2021; Published 13 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Yang 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 establishes a hybrid education teaching practice quality evaluation system in colleges and constructs a hybrid teaching quality evaluation model based on a deep belief network. Karl Pearson correlation coefficient and root mean square error (RMSE) indicators are used to measure the closeness and fluctuation between the effective online teaching quality evaluation results evaluated by this method and the actual teaching quality results. The experimental results show the following: (1) As the number of iterations increases, the fitting error of the DBN model decreases significantly. When the number of iterations reaches 20, the fitting error of the DBN model stabilizes and decreases to below 0.01. The experimental results show that the model used in this method has good learning and training performance, and the fitting error is low. (2) The evaluation correlation coefficients are all greater than 0.85, and the root mean square error of the evaluation is less than 0.45, indicating that the evaluation results of this method are similar to the actual evaluation level and have small errors, which can be effectively applied to online teaching quality evaluation in colleges and universities.

1. Introduction

In the process of informatization teaching with the in-depth integration of information technology and classrooms, teaching methods tend to be more online integrated with offline. At present, we are in a ubiquitous learning environment supported by mobile Internet; how to effectively integrate combining resources, deep learning promoted by design technology, is a question worth exploring and thinking about. With microclasses, MOOCs (massive open online courses), SPOC (small private online course), and flipped classroom development, blended teaching has gradually become one of the important directions of college teaching reform. Blended teaching complements the advantages of traditional classroom teaching and online learning to achieve the optimization of learning effects [1], which is in line with the purpose of deep learning. Through blended teaching, making full use of high-quality resources and tools, and reconstructing the teaching process, an effective teaching plan can be provided for the realization of deep learning.

The current research is mainly concentrated in the field of higher education. The content of the research includes influencing factors, technical support, application models, strategies, blended teaching and online learning, effects and quality, evaluation, and other issues [2]. Based on the analysis of the relevant literature on online learning and blended teaching, Shivetts proposed that students' learning motivation is the main factor for the success of online learning and blended learning. In a blended learning environment, students' learning results and curriculum settings and accessibility sex is closely related [3]. Graff and Ayden conducted research from the perspective of classroom community awareness [4,5], Aspden conducted research from the perspective of student participation and interaction [6], Oliver and Trigwell passed based on the problem-based learning framework to explore the evaluation of blended teaching and so on [7]. Based on the exploration community theory, Garrison et al. further constructed a specific hybrid teaching evaluation framework [8, 9].

The primary task of developing mixed teaching in colleges and universities is how to improve the quality of online teaching and informatization teaching [10]. Therefore, it is very important to study effective methods for evaluating online teaching quality in colleges and universities. Research on online teaching quality evaluation methods in colleges and universities can strengthen online teaching quality management and improve online teaching quality [11]. A deep belief network is an important learning model in deep learning, formed by combining low-level features and more abstract high-level features or attribute categories. It has received widespread attention from all walks of life and has since set off a wave of deep learning research. In recent years, deep belief networks have been widely used in curriculum areas closely related to people's lives and can realize machine translation, face recognition [12], speech recognition, signal recovery, business recommendation, financial analysis, medical assistance, and intelligent transportation. The deep learning network has good robustness and high accuracy. Based on this, this paper studies the method of evaluating the quality of mixed teaching in colleges and universities so as to improve the quality of mixed teaching in colleges.

2. Related Work

2.1. Deep Belief Neural Network. The deep belief network (DBN) model is a series of restricted Boltzmann machine (RBM) models stacked from bottom to top, using an unsupervised greedy layer-by-layer method to pretrain multiple RBMs, which means that each layer of RBM needs to be trained. The weight value enables the hidden layer to obtain the connection of the high-level data expressed by the visible layer [13].

2.1.1. Basic Principles. The training process of the deep belief neural network can be summarized into two parts: one is pretraining, which uses unsupervised layer-by-layer learning to initialize the parameters of the neural network structure, and the other is fine-tuning (fine-training). At the end of the model, the backpropagation algorithm is used to fine-tune the network parameters globally. Even if the entire network has accumulated multiple layers, the parameters can still be reasonably optimized. This learning method solves the problem of gradient disappearance and makes the learning of deep neural networks more efficient.

2.1.2. Restricted Boltzmann Machine. Restricted Boltzmann machine (RBM) originated from the Boltzmann machine (BM). BM not only has a strong unsupervised learning ability but also can learn complex rules in data [14], but this kind of training is relatively complicated, and the training time is longer. In order to overcome this problem, Reikard introduced a restricted Boltzmann machine [15], whose structure is shown in Figure 1.

RBM is composed of a visible layer (visible) and a hidden layer (hidden), and there is no connection in the layer. h_1 to h_n are n real numbers, V_1 to V_m are m real numbers. These

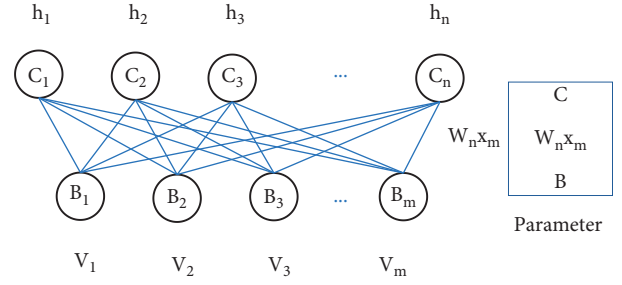


FIGURE 1: RBM model structure diagram.

real numbers are all numbers between 0 and 1, and each of them forms an h vector and a v vector. There is a weight w between each explicit layer and the hidden layer, which has a total weight of n -and m , and c and b are the bias vectors of the hidden and explicit layers, respectively.

RBM is an energy-based model. Then, for a given set of states (v, h) , the energy formula is defined as follows:

$$E(v, h|\theta) = -\left(\sum_{ij} w_{ij} v_i h_j + \sum_i b_i v_i + \sum_j c_j h_j\right). \quad (1)$$

Among them, θ is the parameters w , c , and b . The energy represented on the right side of the equation has three parts. One is generated by the weight w connecting the nodes v and h on both sides, and all three of them must be 1 to be considered as energy output; the other two are the multiplication of the offset on the node and the vector dimension value of the node input, and the same three must be 1 to be considered as energy output. Each energy corresponds to a state; the smaller the energy, the more stable the model. When the parameters are determined, the joint probability distribution of the explicit layer and hidden layer neurons can be obtained from the energy formula as follows:

$$P_{\theta}(v, h) = \frac{1}{Z(\theta)} e^{-E(v, h|\theta)}, \quad (2)$$

$$Z(\theta) = \sum_{v, h} e^{-E(v, h|\theta)}. \quad (3)$$

Therefore, when determining the state of the input layer, the activation states of the hidden layer nodes are independent of each other. Therefore, the activation probability of the i -th hidden layer node is as follows:

$$P(h_i = 1|v) = \sigma\left(\sum_{j=1}^m w_{ij} \times v_j + c_i\right), \quad (4)$$

$\sigma(x) = 1/1 + \exp(-x)$ is the activation function (Sigmoid).

Given that the structure of the restricted Boltzmann machine is symmetrical, when determining the state of the hidden layer node, the activation state of each input layer node is also conditionally independent, so in the same way, the activation probability of the j -th visible unit can be expressed as follows:

$$P(v_i = 1|h) = \sigma \left(\sum_{j=1}^n w_{ij} \times h_j + b_i \right). \quad (5)$$

2.1.3. Algorithm Execution Process of the Restricted Boltzmann Machine. In 2002, Brosch proposed a fast-learning algorithm for RBM-Contrastive Divergence (CD) [16]. In the CD algorithm, the state of the visible layer node can be used as the data feature value, and the activation state of all hidden layer nodes is calculated by formula (4). After the states of all hidden layer nodes are determined, the probability is calculated that the value of the i -th input layer node is 1 according to formula (5), and then, a reconstruction of the visible layer is obtained. In this way, when using the value of the log-likelihood function of the stochastic gradient ascent method on the training data, the update criterion of each parameter is as follows:

$$\Delta W_{ij} = e \left(\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}} \right), \quad (6)$$

$$\Delta b_i = e \left(\langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{recon}} \right), \quad (7)$$

$$\Delta b_i = e \left(\langle h_i \rangle_{\text{data}} - \langle h_i \rangle_{\text{recon}} \right). \quad (8)$$

Among them, e is the learning rate, $\langle h_i \rangle_{\text{data}}$ is the mathematical expectation defined by the training dataset, and $\langle h_i \rangle_{\text{recon}}$ represents the expectation of the model definition after a step of reconstruction.

2.1.4. Basic Structure. The deep belief network is composed of multiple restricted Boltzmann machines and a layer of supervised classifiers, as shown in Figure 2. DBN can be summarized into two stages in the whole learning process. the first stage is unsupervised learning, and the second stage is supervised fine-tuning [17, 18].

- (1) The first stage is to conduct greedy training on multilayer RBM, that is, training layer by layer, and training the model layer by layer from the bottom to the top. The output value of the previous RBM is used as the input value of the next RBM to realize the initialization of the network parameters.
- (2) The second stage is to fine-tune the model. In order to optimize the objective function of the entire model structure, the BP neural network algorithm or the support vector machine can be used to fine-tune the parameters to achieve global optimization. The fine-tuned network parameters are used as the initial parameters of the entire network. Compared with traditional neural networks, DBN has a higher accuracy rate and at the same time solves the problem of easily falling into local optimality.

2.2. Mixed Education Evaluation System. Blended teaching is a new type of education and teaching mode that combines the advantages of traditional face-to-face teaching and online teaching, and integrates information

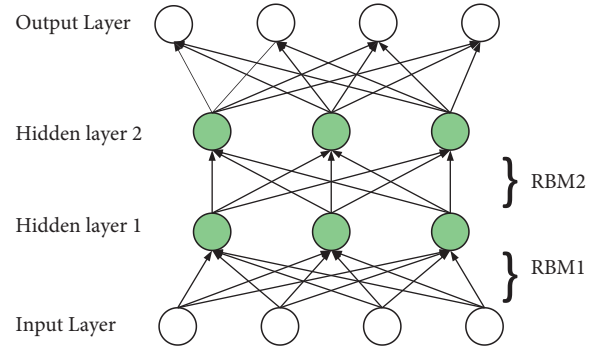


FIGURE 2: The structure of DBN model.

technology with traditional education and teaching in an effective form. The realization of this comprehensive advantage is based on the deep integration of different application methods, different teaching concepts, and different technical means. Blended education teaching has five basic characteristics: (1) the integration of teaching theories; (2) the integration of teaching resources; (3) the integration of teaching environments; (4) the integration of teaching methods; (5) the integration of teaching strategies; and (6) the integration of evaluation methods [19].

Through reading related literature, the teaching evaluation camp covers the entire process of teacher teaching, basically including teaching objectives, teaching design, teaching process, and teaching effect. The concept of “assessing teaching by learning” and the characteristics of a mixed teaching model are constructed in this paper to construct a three-level evaluation index. In order to maintain the consistency of descriptions among indicators at all levels, this article describes the first-level indicators of mixed teaching evaluation in terms of “learning design,” “learning environment,” “learning process,” and “learning effect.”

the above-mentioned first-level index system is used to expand the hybrid education and the teaching evaluation system, the hybrid education and teaching evaluation indicators are expanded, and a hybrid education and the teaching evaluation system are built. The specific evaluation methods are shown in Table 1.

3. The Quality Evaluation Model of Hybrid Higher Education Teaching Practice Based on Deep Learning Route

3.1. Model Structure. DBN is a generative model based on unsupervised learning, and most of the objects it faces are unknown data, such as nonlinear systems. The main goal of deep learning is to use algorithms to describe changes in data [18]. However, most DBNs build models with dense representations, and there is a big problem with this representation; that is, any fluctuations or noise in the input will greatly change the feature representation vector extracted by the hidden layer, which will lead to the robustness of the network. The model is poor [21].

TABLE 1: Hybrid higher education teaching quality evaluation system [20].

Evaluation object	First-level index	Second-level index	Third-level index
Blended teaching quality	A1 learning design	B1 learning objectives	C1 learning objectives at different levels of knowledge, abilities, qualities, etc., are positioned accurately and comprehensively, and the online and offline learning objectives are distinct
		B2 learning strategy	C2 learning strategy matches with learning goals C3 learning strategies are in line with students' academic conditions C4 learning strategies are to meet the learning needs of students
		B3 learning methods	C5 learning methods are diverse and interactive
		B4 learning support	C6 learning platform provides guarantee in terms of space, time, and equipment
	A2 learning environment	B5 learning resources	C7 online learning resources are highly learnable and pertinent C8 offline learning resources are more thoughtful and inquiring C9 expansion of learning resources is hierarchical and challenging
		B6 learning links	C10 online and offline learning links are reasonably distributed C11 online and offline learning links are closely connected C12 learning content is highly learnable
	A3 learning process	B7 learning content	C13 online and offline learning content complements C14 learning content is cutting-edge and challenging
		B8 learning participation	C15 Student's investment in various teaching links online and offline
		B9 learning assessment	C16 students' thinking and feedback on C17 online learning assessment has clear levels
		B10 learning willingness	C18 offline classroom learning assessment is innovative and challenging C19 increased students' learning enthusiasm and sense of learning achievement
	A4 learning effect	B11 learning ability	C20 students' enthusiasm and initiative to explore challenging learning content C21 increased willingness of students to display their personal learning achievements C22 improve students' autonomous learning ability
		B12 learning quality	C23 students perform well in case analysis, experimental operation, and situational practice C24 students can take the initiative to ask questions during the teaching process and propose solutions to the knowledge they have learned
		C25, the improvement degree of students' knowledge, ability, quality, and ability matches the learning goals, the goal achievement degree is high, and the students are highly satisfied with the course teaching	

This paper introduces an output layer in the top layer of the deep belief network model and implements score level mapping to the characteristics of all colleges and universities' online teaching quality evaluation indicators through the Softmax multiclass classifier based on logistic regression expansion and completes the online teaching quality evaluation of colleges and universities [23]. The online teaching quality score mapping of colleges and universities is a multiclass classification problem, so the extended logistic regression is a Softmax regression. Let $\{(x_1, y_1), \dots, (x_n, y_n)\}$ be the set of the state x of each neuron in the hidden layer and the corresponding class label y of each neuron; for M categories, the corresponding class of the i -th feature for labels $y_i \in \{1, 2, \dots, M\}$, the M -dimensional category probability distribution matrix is as follows:

$$l_\sigma(x^i) = \begin{pmatrix} p(y^i = 1|x^i; \sigma) \\ p(y^i = 2|x^i; \sigma) \\ \dots \\ p(y^i = m|x^i; \sigma) \end{pmatrix} = \frac{1}{\sum_{j=1}^M e^{\sigma_j^T x^i}} \begin{pmatrix} \sigma_1^T x^i \\ \sigma_2^T x^i \\ \dots \\ \sigma_M^T x^i \end{pmatrix}, \quad (9)$$

$$\sigma = \begin{pmatrix} \sigma_1^T \\ \sigma_2^T \\ \dots \\ \sigma_M^T \end{pmatrix}.$$

In the above formula, the model parameter matrix of $M \times N$ is σ ; the input feature dimension is N . The expression formula of the cost function of Softmax regression is as follows:

$$J(\sigma) = \frac{1}{n} \left[\sum_{i=1}^n \sum_{j=1}^M 1\{y^i = j\} \log \frac{e^{\sigma_j^T x^i}}{\sum_{d=1}^M e^{\sigma_d^T x^i}} \right] = \frac{\tau}{2} \sum_{i=1}^M \sum_{j=1}^N \sigma_{ij}^2. \quad (10)$$

The college hybrid teaching quality evaluation system is used as the input value of the deep belief network's online teaching quality evaluation model of colleges and universities. To ensure that the deep belief network model conforms to the characteristics of the input college online teaching quality evaluation system, the first layer of RBM uses Gauss-Bernoulli. RBM is used to map the features entered into the online teaching quality evaluation system of colleges and universities into a binary state. The remaining layers of RBM use Bernoulli-Bernoulli RBM, and the Bernoulli-Bernoulli RBM abstracts the features between the essential relationship. The hybrid higher education teaching quality evaluation model based on the DBN model is shown in Figure 3.

3.2. Training Steps. Since the $p(h_1; W_1)$ of the deep belief network model is consistent with the $p(v|h_1; W_1)$ of the RBM, there is a greedy layer-by-layer pretraining step of the s-layer hidden deep belief network model in the model.

Step 1. Train RBM to obtain the weight parameter W_1 of the bottom RBM.

Step 2: Start training the second layer of RBM, initialize the second layer weight parameter $W_2 = W_1 T$ to ensure that the deep belief network model with two hidden layers is better than the uninitialized RBM, and then pass fixed W_1 to train the second layer of RBM to optimize W_2 .

Step 3: Continue to train the third layer of RBM, initialize the third layer RBM weight parameter $W_3 = W_2 T$, train the third layer RBM by fixing W_2 to make W_3 reach the optimal value, similarly, complete the final layer of RBM weight to reach the optimal value, complete the unsupervised pretraining of the entire deep belief network model.

In the case of optimal learning of the parameters to be tested, for the feature to be tested, and after learning the deep belief network model of the trained model parameters, the probability that this feature belongs to the online teaching quality scores of various colleges and universities is calculated, and the quality score with the highest probability is selected as output result.

4. Simulation Experiment

4.1. Experimental Data. Taking a university as the evaluation object, the mixed teaching data of the university from March to August of 2020 are randomly selected as the dataset. The data include a total of 5,000 data samples. The online teaching quality evaluation method for colleges and

universities based on the deep learning network studied in this paper is used to implement online teaching quality evaluation. The selected evaluation indicators are shown in Table 2. The scoring method of the evaluation indicators is to select 10 experts from outside the school and 5 senior professors familiar with the online teaching situation in the school. The average scoring value is regarded as the final evaluation score of each secondary evaluation index. The interval of this value is (0–1). Each score corresponds to the online teaching quality evaluation level of colleges and universities; that is, the evaluation criteria are shown in Table 3.

4.2. Learning and Training Performance. In order to test the learning performance and training performance of the DBN model used in this paper, 3,000 data samples out of 5,000 data samples are used as the model training set. The remaining 2,000 data samples are used as the model test set, and the training set sample data are used as the model input, the model training is expanded, and the model fitting error is calculated under different iteration times in the training process. The result is shown in Figure 4. It can be seen from the data in Figure 1 that as the number of iterations increases, the fitting error of the DBN model used in the method in this paper decreases significantly. When the number of iterations reaches 20, the fitting error of the DBN model stabilizes and decreases to below 0.01. The experimental results show that the model used in this method has good learning and training performance, and the fitting error is low.

The 2,000 test samples in the test set are divided into 10 groups and input them into the model for model testing, and the output results of the model are compared with the expected output results. The results are shown in Table 4. From the data in Table 4, it can be seen that the relative error between the test output result and the expected output result after the model is trained is less than 3%, and the evaluation level of the method in this paper is exactly the same as the actual level. Experimental results show that the trained model has evaluation accuracy and can be effectively used for online teaching quality evaluation in colleges and universities.

4.3. Evaluate Performance. In order to verify the evaluation performance of the method in this paper, Karl Pearson correlation coefficient and RMSE are used to measure the degree of closeness and fluctuation between the evaluation results of efficient online teaching quality evaluated by this method and the actual teaching quality results. The correlation coefficient R and the root mean square error RMSE are counted in the evaluation process of the method in this paper, and the results are shown in Table 5. Analyzing the data in Table 5, we can see that the evaluation correlation coefficients of the method in this paper are all greater than 0.85, and the root mean square error of the evaluation is less than 0.45. The results show that the evaluation results of the method in this paper are highly similar to the actual evaluation level, and the error is small, which can be effectively

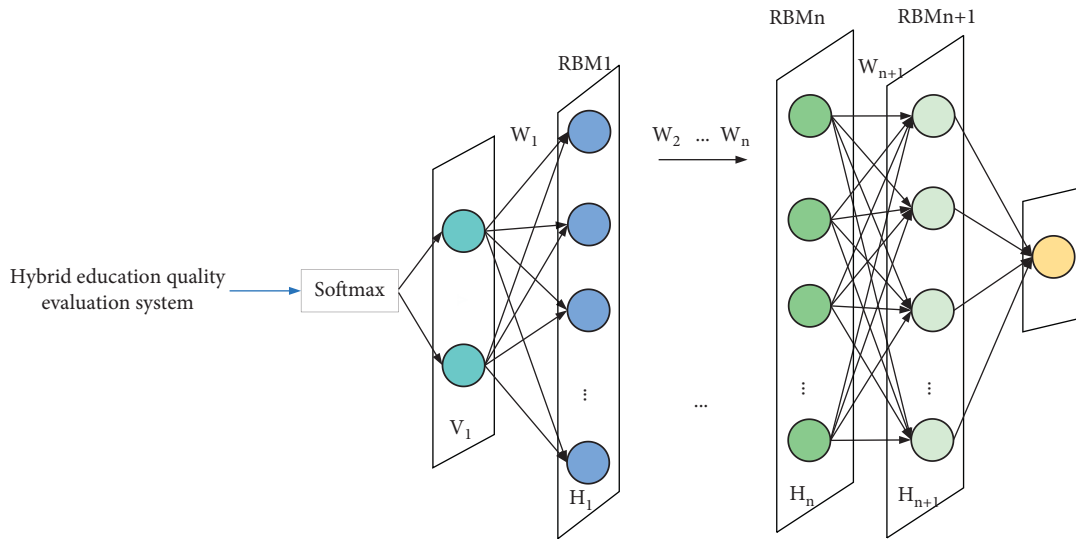


FIGURE 3: A hybrid university teaching quality evaluation model based on the deep belief network.

TABLE 2: Blended education and teaching evaluation system-first level.

Indicators	Index definition
Learning design	The learning design index refers to the preparatory part of the teacher’s preparation for student learning in the mixed teaching process. It guides the development of the entire teaching activity and is the basis of the entire teaching activity. Therefore, it is reflected from the three aspects of learning objectives, learning strategies, and learning methods.
Learning environment	The learning environment index refers to the physical environment in the mixed teaching mode, which provides support for the development of mixed teaching. The influence of this kind of learning environment on the quality and effect of blended teaching cannot be ignored. Therefore, it is reflected from the two aspects of learning support and learning resources.
Learning process	The learning process indicators cover teacher teaching and student learning content, behavior, and assessment in each link of blended teaching. Combining the characteristics of the blended teaching model, this indicator reflects the four aspects of learning links, learning content, learning participation, and learning assessment.
Learning effect	The learning effect index includes two aspects: student learning and teacher teaching. Student learning is reflected in two aspects: willingness to learn and learning ability while teacher teaching is reflected in learning quality.

TABLE 3: Hybrid higher education teaching quality evaluation standard.

Evaluation score	Estimated grade
>0.75	Excellent
0.5~0.75	Good
0.25~0.49	Normal
<0.25	Bad

applied to online teaching quality evaluation in colleges and universities.

Through the above experiments, it can be seen that the method in this paper has good performance in evaluating the quality of online teaching in colleges and universities. For this reason, the method in this paper is used to evaluate the online teaching quality of the colleges and universities. The evaluation results are shown in Table 6. After analyzing the data in Table 6, it can be seen that the method in this paper can realize the quality evaluation of the mixed education of the university by evaluating the various indicators that measure the quality of the mixed education of the university.

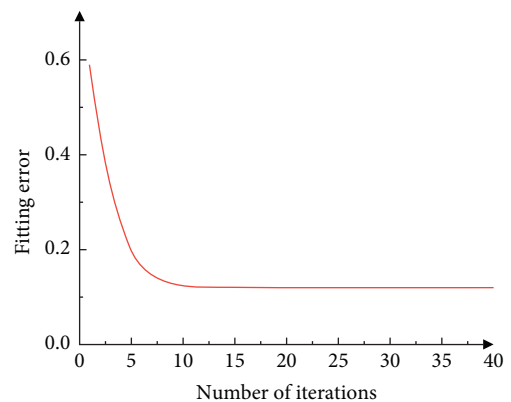


FIGURE 4: The learning performance and training performance of the DBN model.

The mixed education quality of the college is generally evaluated by the method of this article. Colleges and universities can take targeted measures based on the indicators with lower scores.

TABLE 4: Model test results.

Dataset	Expected output	Actual grade	Test output	Evaluation level	Relative error (%)
1	0.78	Excellent	0.79	Excellent	1.28
2	0.59	Good	0.58	Good	1.7
3	0.42	Normal	0.43	Normal	2.38
4	0.52	Good	0.53	Good	1.92
5	0.11	Bad	0.11	Bad	0
6	0.24	Bad	0.24	Bad	0
7	0.88	Excellent	0.87	Excellent	1.14
8	0.46	Normal	0.47	Normal	2.17
9	0.82	Excellent	0.81	Excellent	1.22
10	0.65	Good	0.66	Good	1.54

TABLE 5: R and RMSE statistical results.

Dataset	R	RMSE
1	0.92	0.25
2	0.88	0.33
3	0.95	0.41
4	0.89	0.45
5	0.92	0.28
6	0.95	0.17
7	0.89	0.41
8	0.86	0.22
9	0.87	0.34
10	0.96	0.26

TABLE 6: Quality evaluation results of higher mixed education.

First-level index		Second-level index		Evaluation result
Number	Weight	Number	Weight	
A1	0.2	B1	0.4	0.75
		B2	0.5	0.85
		B3	0.1	0.74
A2	0.2	B4	0.5	0.79
		B5	0.5	0.72
		B6	0.2	0.56
A3	0.4	B7	0.3	0.71
		B8	0.3	0.69
		B9	0.2	0.75
A4	0.2	B10	0.3	0.46
		B11	0.4	0.62
		B12	0.3	0.71

The evaluation results of this article is to improve the quality of online teaching.

5. Conclusion

This paper establishes a hybrid education teaching practice quality evaluation system in colleges and universities, and constructs a hybrid teaching quality evaluation model based on a deep belief network. Karl Pearson correlation coefficient and Root Mean Square Error (RMSE) indicators are used to measure the closeness and fluctuation between the effective online teaching quality evaluation results evaluated by this method and the actual teaching quality results degree. The experimental results show that:

- (1) As the number of iterations increases, the fitting error of the DBN model decreases significantly. When the number of iterations reaches 20, the fitting error of the DBN model stabilizes and decreases to below 0.01. The experimental results show that the model used in this method has good learning and training performance, and the fitting error is low.
- (2) The evaluation correlation coefficients are all greater than 0.85, and the root mean square error of the evaluation is less than 0.45, indicating that the evaluation results of this method are similar to the actual evaluation level and have small errors, which can be effectively applied to online teaching quality evaluation in colleges and universities.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Y. Rizk, N. Hajj, and N. Mitri, "Deep belief networks and cortical algorithms: a comparative study for supervised classification," *Applied Computing and Informatics*, vol. 15, pp. 81–93, 2019.
- [2] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural Network Design*, PWS Publishing, Boston, MA, USA, 1996.
- [3] C. Shivettis, "E-learning and blended learning: the importance of the learner—a research literature review," *International Journal on E-Learning*, vol. 10, no. 3, pp. 331–337, 2011.
- [4] M. Graff, "Individual differences in sense of classroom community in a blended learning environment," *Journal of Educational Media*, vol. 3, pp. 203–210, 2003.
- [5] E. Ayden and S. Gums, "Sense of classroom community and team development process in online learning," *Turkish Online Journal of Distance Education*, vol. 17, 2013.
- [6] L. Aspden and P. Helm, "Making the connection in a blended learning environment," *Educational Media International*, vol. 41, no. 3, pp. 245–252, 2004.
- [7] M. Oliver and K. Trigwell, "Can blended learning Be redeemed?" *E-learning*, vol. 2, no. 1, pp. 17–26, 2005.
- [8] D. R. Garrison, "Online community of inquiry review: social, cognitive, and teaching presence issues[J]," *Journal of Asynchronous Learning Networks*, vol. 11, no. 1, pp. 61–72, 2007.

- [9] D. R. Garrison, M. Cleveland-Innes, and T. S. Fung, "Exploring causal relationship among teaching, cognitive and social presence: student perceptions of the community of inquiry framework," *Internet & Higher Education*, vol. 23, no. 1-2, pp. 31–36, 2010.
- [10] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.
- [11] A. Ali and F. Yangyu, "Automatic modulation classification using deep learning based on sparse autoencoders with nonnegativity constraints," *IEEE Signal Processing Letters*, vol. 24, no. 11, pp. 1626–1630, 2017.
- [12] S. Jagannathan and F. L. Lewis, "Identification of nonlinear dynamical systems using multilayered neural networks," *Automatica*, vol. 32, no. 12, pp. 1707–1712, 1996.
- [13] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009.
- [14] G. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 1989.
- [15] G. Reikard, "Predicting solar radiation at high resolutions: a comparison of time series forecasts," *Solar Energy*, vol. 83, no. 3, pp. 342–349, 2009.
- [16] T. Brosch, Y. Yoo, and D. K. B Li, "Modeling the variability in brain morphology and lesion distribution in multiple sclerosis by deep learning," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 462–469, Boston, MA, USA, September 2014.
- [17] S. Kamada, T. Ichimura, and T. Harada, "Adaptive structural learning of deep belief network for medical examination data and its knowledge extraction by using C4.5," in *Proceedings of the IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, pp. 33–40, Laguna Hills , CA, USA, 2018.
- [18] I. A. E. Spanjers, "The promised land of blended learning: quizzes as a moderator," *Educational Research Review*, vol. 15, pp. 59–74, 2015.
- [19] A. Naylor and J. Gibbs, "Deep learning," *International Journal of Mobile and Blended Learning*, vol. 10, no. 1, pp. 62–77, 2018.
- [20] M. B. Horn and H. Staker, *Blended Using Disruptive Innovation to Improve schools*, Wiley, Hoboken, New Jersey, USA, 2015.
- [21] M. A. Schnabel and J. J. Ham, "The social network virtual design studio: integrated design learning using blended learning environments," *Architectural Design*, vol. 83, no. 2, pp. 118–123, 2013.

Retraction

Retracted: Realization of Super-Large-Diameter Slurry Shield Passing through Settlement-Sensitive Area Based on Unreinforced Disturbance Control Technology

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. 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 article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

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

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

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

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

References

- [1] D. Liu, X. Liu, Z. Zhong, Y. Han, F. Xiong, and X. Zhou, "Realization of Super-Large-Diameter Slurry Shield Passing through Settlement-Sensitive Area Based on Unreinforced Disturbance Control Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6299645, 9 pages, 2022.

Research Article

Realization of Super-Large-Diameter Slurry Shield Passing through Settlement-Sensitive Area Based on Unreinforced Disturbance Control Technology

Dongshuang Liu,¹ Xinrong Liu ,^{1,2,3} Zuliang Zhong ,^{1,2,3} Yafeng Han,¹ Fei Xiong,¹ and Xiaohan Zhou ¹

¹School of Civil Engineering, Chongqing University, Chongqing 400045, China

²National Joint Engineering Research Center of Geohazards Prevention in the Reservoir Areas (Chongqing) of Chongqing University, Chongqing 400045, China

³Key Laboratory of New Technology for Construction of Cities in Mountain Rrer of Chongqing University, Chongqing 400045, China

Correspondence should be addressed to Xinrong Liu; liuxr666@126.com

Received 20 November 2021; Revised 16 December 2021; Accepted 24 December 2021; Published 12 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Dongshuang 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.

Due to the complex construction conditions of shield tunnels, ground disturbance is inevitable during the construction process, which leads to surface settlement and, in serious cases, damage to surrounding buildings (structures). Therefore, it is especially important to effectively control the constructive settlement of subway tunnels when crossing settlement-sensitive areas such as high-density shantytowns. Based on the project of Wuhan Metro Line 8 Phase I, the shield of Huangpu Road Station-Xujiapang Road Station interval crossing high-density shantytowns, we study the disturbance control technology of oversized diameter mud and water shield crossing unreinforced settlement-sensitive areas during the construction process. By optimizing the excavation parameters and evaluating the ground buildings, the excavation process can be monitored at the same time, and the water pressure, speed, and tool torque required during the excavation during the construction process can be finely adjusted; the control of tunneling process parameters can provide reference and basis for analyzing the construction control of large-diameter shield through old shantytowns.

1. Introduction

With the rapid development of urban rail transit construction, complex situations such as mud-water balanced shield crossing complex strata and dense old building complexes are becoming more and more frequent [1–3]. Around the settlement control of tunnel shield under old buildings, many useful studies have been done by domestic and foreign research scholars in different engineering contexts [4]. After a long investigation on the stability and settlement law of the upper buildings during the shield underpass, it is concluded that in general, the surface settlement of the upper buildings

accounts for 30% to 90% of the total amount during the shield underpass. Wang et al. [5] studied the influence of the tunnel shield excavation process on the settlement of the upper structures under different burial conditions. Charaniya, Nassar, and Al-Mahdi [6, 7] studied the problem of surface settlement and deformation; that is, when the actual site water pressure is high, the water storage tunnel encounters soil particles with strong water permeability. Rho et al. and Reed et al. [8, 9] simplified the calculation method for such problems and simplified the complex three-dimensional problem into a planar problem, which greatly reduced the calculation volume and accelerated the calculation process.

People's living space is getting more and more crowded, and traffic congestion is becoming more and more prominent. The rational exploitation of underground space can greatly help to relieve the pressure of urban traffic, and the development of underground rail transportation has become one of the important directions for sustainable social development [4, 10]. China is developing rapidly in urban rail transit construction and currently ranks first in the world in total subway construction.

In urban metro construction, tunnel construction mainly includes the open cut method, shield method, and mining method. Due to the advantages of short construction time, high engineering safety, high automation, and low impact on the surrounding environment [11], shield method construction has now gradually replaced other methods as the first choice for urban metro construction. The shield method construction process is a process in which the shield machine excavates and digs forward in the underground soil layer, and the shield shell and tube sheet are used to maintain the stability of the soil around the excavation and avoid excessive soil deformation, while the soil is excavated by the cutting device in front of the excavation, and the cut down soil is cut by the shield machine [12]. The excavation is carried out by the cutting device in front of the excavation, and the cut soil is transported out of the hole by the excavation machine inside the shield machine. The jack at the end of the shield acts on the completed tube sheet, and the jack pushes the shield forward while the tube sheet is released from the end of the shield to form the tunnel structure. During the tunneling process, as the diameter of the cutter plate and the outer diameter of the shield excavated in front of the shield are larger than the outer diameter of the tunneled pipe pieces inside the shield, when the pipe pieces are assembled and completely disengaged from the tail of the shield machine, an annular gap will be formed between the pipe pieces and the soil layer, which is called the shield tail gap [13].

Usually, post-wall grouting is used to fill the void at the end of the shield, and the post-wall grouting construction technology is simple and has good engineering effect. Therefore, it is widely used in tunneling and other geotechnical engineering. Post-wall grouting is relatively mature; however, the movement of the slurry in the shield tail void is complicated. There is still a lot of disagreement in the form of pressure distribution for post-wall grouting in shield tunnels. There are still many differences in the form of pressure distribution, and the mechanism of slurry action with the surrounding strata and tube sheet is still to be investigated. Therefore, the mechanism of post-wall grouting has been a topic of interest for many researchers and scholars [8, 14]. The current grouting studies are basically for small and medium-sized shield structures, and the effect of gravity on the grouting pressure is not considered, but the effect of gravity on the grouting pressure of large-diameter shield structures cannot be ignored. Therefore, it is necessary to analyze the construction process of post-wall grouting of large-diameter mud and water shield [15].

This study takes the Wuhan Rail Transit Line 8 Huangpu Road Station-Xujiapang Road Station river-crossing shield tunnel project as an example and proposes construction

control techniques for shielding crossing-sensitive areas in dense shantytowns based on a brief analysis of the project background and key difficulties.

2. Related Work

The history of behind-the-wall grouting was developed as part of geotechnical grouting for shield tunneling applications, which opened the door to the use of grouting in geotechnical engineering in 1802 when Charles Berigny, a French civil engineer, used manual hammering to inject clay and lime into the ground by means of a wooden impact cylinder device when dealing with sand gates. Zaslow et al. [16] from the Research Institute of China Railway Tunneling Group developed a new type of single-liquid slurry, which has a long setting time of more than 3–12 h and can be adjusted to adapt to different geological conditions and construction situations. By selecting different water-reducing agents for the test, Qiang et al. [17] improved the pumpability of the slurry, improved the stability of the slurry, improved the construction performance of the post-wall synchronous injection slurry, and obtained a good grouting effect.

In the tunneling project of Beijing Metro Line 5, Zhou et al. [18] studied the composition of the slurry, and finally the inert slurry was selected for the slurry used in the construction, which is based on quicklime and fly ash as the main ingredients, and a patent was published for this slurry. The slurry has the characteristics of low manufacturing cost, simple ingredients, good compatibility, and easy production and also has the advantages of preventing leakage from the longitudinal joints of the pipe sheet ring, convenient construction management, and less ground settlement. Liu et al. [19] conducted a lot of research on the admixtures of post-wall grouting materials. Usually, synchronous grouting materials contain bentonite compounded with water-reducing agent, and Li et al. [20] found that the grouting effect is better with the choice of water-retaining agent compared to water-reducing agent, and the cost of both is similar, so the water retaining agent is worth using widely. For post-wall slurry material delivery, Lai et al. [21] invented modified lignin efficient grouting pumping agent which can improve the stability and water-reducing enhancement of slurry while reducing the pumping resistance of slurry and improving its strength, so that the slurry has a proper setting time.

In the South-North Water Transfer Central Line crossing Yellow River Tunnel Project, Xie et al. [22] studied how to optimize the ratio of grouting materials based on the experimental principle of uniform design for fine sandy soils and compared and analyzed the slurry performance influenced by each component of the single-liquid active grouting material.

For the effect of post-wall grouting, the first research in this area was carried out and some results were obtained by the whole group led by Professor Huang Hongwei from Tongji University. In a subway shield tunnel project in Shanghai, Qiu et al. [23] studied the deformation properties and longitudinal mechanics of shield tunnels, paying special

attention to the effect of behind-wall synchronous grouting, and for the first time used ground-penetrating radar method to detect the distribution of tunnel synchronous grouting slurry, proposing that the form of behind-wall slurry distribution of the tube piece can be effectively detected by this method, which provides the control of longitudinal mechanical deformation properties of tunnel grouting, laying the foundation for the longitudinal mechanical deformation control of tunnel grouting.

Field measurements of the tube lining pressure revealed that the dissipation time of the grouting pressure has a significant effect on the variation of the tube lining pressure. For 9.5 m diameter shield tunnels, Huang et al. [24] conducted a field experimental study on simultaneous post-wall grouting. After analyzing the field monitoring data, it was found that the surface settlement was significantly affected by the grouting pressure, and the distribution of the grouting pressure in the cohesive soil was found to be small at the top and large at the bottom when the shield was stopped. In order to achieve accurate measurement of the pressure effect of post-wall grouting on the pipe sheet lining, Yin et al. [25] developed a gasket type lining pressure measurement meter, which is capable of long-term uninterrupted monitoring of the pipe sheet pressure as well as the pressure of the grouting layer. Analysis of the field monitoring data revealed that the main reason for the change in pipe sheet pressure was the change in grouting pressure. Meschke [26] took the surrounding rock of a project shield tunnel as the object of study and used field monitoring method for this surrounding rock pressure, illustrating the selection and burial of measuring elements required for monitoring, testing methods, and analysis of test results. Li and Yuan [27] used monitoring instruments to measure the dry shrinkage stresses applied to the tube sheet lining, the post-wall grouting tube sheet lining stresses, the tube sheet lining stresses after stabilization of the grouting ring, and the surrounding rock pressure in the shield tunnel during construction. The corresponding analytical solutions were obtained by using the steel stress back calculation formula combined with the monitoring data for the annular pressure of the pipe sheet and the bending moment. In [28], the interaction between the tube sheet and the surrounding soil in a shield tunnel through old loess was studied by means of real-time on-site tracking and monitoring.

Through a model test study of post-wall grouting, Zhao et al. [29] found that tunnel pressure loads and soil displacements were significantly influenced by grouting pressure and soil density. In the centrifugal model test, Meng et al. [30] used a mini-shield machine to simulate the shield boring process, including the release of the tube sheet lining from the shield tail, for shield tunnels with different burial depths with or without shield tail voids and derived the deformation of the surrounding strata and the variation of the soil pressure on the top of the tube sheet. Zhang et al. [31] conducted an experimental study using a homemade behind-the-wall grouting test device to study the deformation of the grouted body and the variation of

the grouting pressure considering the grouting material, the grouting pressure, and the geological conditions and concluded that the surface settlement was significantly affected by the grouting pressure and the geological conditions. Many physical model tests of post-wall grouting were conducted by Merceron and Prouteau [3], and it was found that the soil pressure distribution was significantly influenced by the grouting pressure and soil compaction. The shield tunneling in loose sand and dense sand layers was studied separately.

In [4], the slurry diffusion radius and the pressure applied to the pipe sheet were calculated for the slurry under two conditions, considering viscosity and not considering viscosity, respectively. Wang et al. [5] established a three-dimensional finite element model for mud-water balanced shield tunneling, in which the soil was simulated by elastomer, the lining was simulated by shell unit, and the slurry injection was simulated by solid unit, and the support pressure was applied at the excavation surface during shield tunneling, and the slurry injection pressure was considered at the shield tail gap, and the effects of the support pressure and slurry injection pressure on the ground displacement were analyzed.

3. Project Overview

Located 450 m upstream of Wuhan Yangtze River Second Bridge, the Wuhan Rail Transit Line No. 8 Phase I civil construction tunnel project, Huangpu Road Station-Xujiapang Road Station, is a 12.1 m diameter single tube double line composite lining shield tunnel. The shield interval is 3.185.5 m long. After the longitudinal section of the river-crossing line leaves Huangpu Road Station, the line descends with a longitudinal slope of -18.4% and a slope length of 1,710 m. It reaches the lowest point of the river line with a gentle slope of $+4.8\%$ and a slope length of 420 m. It then ascends with a slope of $+27.49\%$ and a slope length of 952.708 m to Xujiapang Road Station [14].

The shield traverses the stratum as shown in Figure 1. The shield traverses 1.820 m of full-section fine sand strata and 1.365 m of soft and hard composite strata with a maximum compressive strength of 24.5 MPa of cemented rocks.

The shield tunnel starts from Xujiapang Road Station with a radius curve of $R = 700$ m and crosses a high-density shantytown (DK12 + 185.000-DK12 + 939.618, total length of 754.618 m) for a long distance, mainly passing through buildings (structures) such as Wumen Village, Station New Village, Xujiapang Street, Chengsanli, and other villages with 1 to 4 floors, with a total of 140 houses passing underneath and 308 houses passing sideways, part of which is shown in Figure 2. Most of these houses were built before the 1990s, and the appearance of these buildings is old and obsolete, and most of them have to peel walls and cracked walls, which are serious safety hazards. Figure 2 shows the plan of Huangpu Road Station-Xujiapang Road Station through some shantytowns.

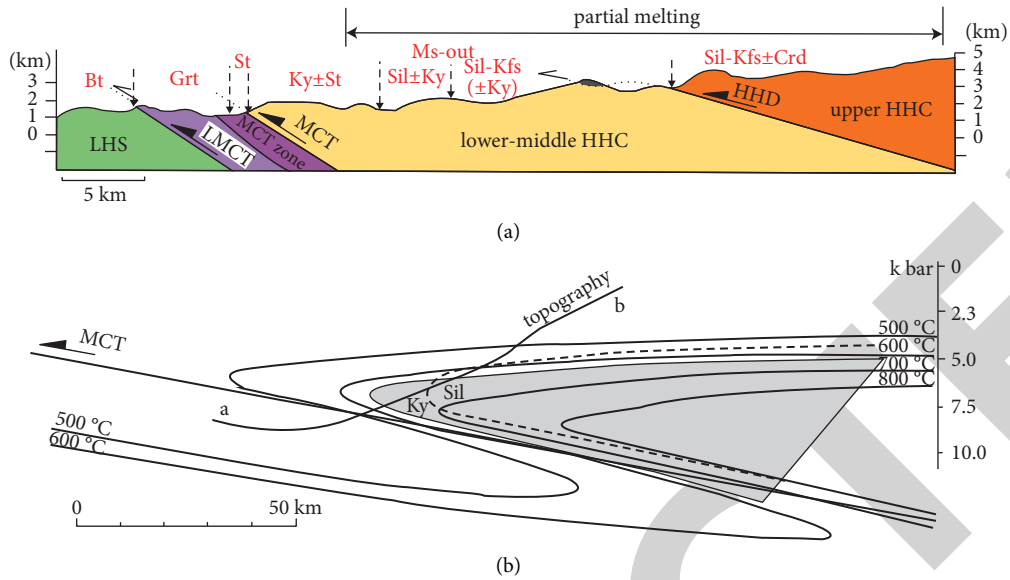


FIGURE 1: Shield crossing geological section.



FIGURE 2: Tunnel and shantytown flat location map.

4. Construction Control Key Technology

4.1. *Pre-Digging Settlement Control Technology.* Before the assessment of existing buildings and safety standards were formulated. Relevant units were commissioned to conduct a safety risk assessment of existing buildings along the shantytown. This interval has 2 level-1 environmental risk projects, and the main measures for shield construction through the level-1 risk project profile are as follows: 15 m range on each side of the tunnel centerline over the risk source is the strongly affected area, and the residents in the houses in the area are to be transitioned before the construction of the interval; the interval across the river crosses the railroad track and cross-line flyover at Wuchang North Station in Wuchang section, and pre-reinforcement measures (cuff tube grouting, rotary piles, and three-row composite anchor piles) are taken before the construction for the construction of the interval.

The initial 75 m of shield advance was used as the test section, and the shield boring parameters were continuously

adjusted during the shield boring process based on real-time monitoring. The construction parameters of the shield in the section under the old shantytown are shown in Table 1.

During the shield tunneling process, the shield machine controls the tunneling attitude and adjusts each tunneling parameter in real time according to the generated ground settlement to ensure the safety and stability of the surrounding soil during the construction process, for example, the relationship between the pressure difference between the inside and outside of the soil layer and the ground settlement is shown in Figure 3.

5. Analysis of Digging Parameters

The geological condition and surrounding environment are important parameters in the process of shield tunneling, and the parameters are set reasonably by combining relevant information. During excavation, we monitor and record the shield operation and changes in excavation parameters and

TABLE 1: Shield tunneling-related technical indicators.

Shield index	Numerical value
Maximum thrust	278400 KN
Normal thrust	180038 KN
Maximum propulsion speed	50 mm/min
Rated torque of cutter head	36585 kN m
Maximum torque of cutterhead	43902 kN m
Cutter head release torque	54878 kN m

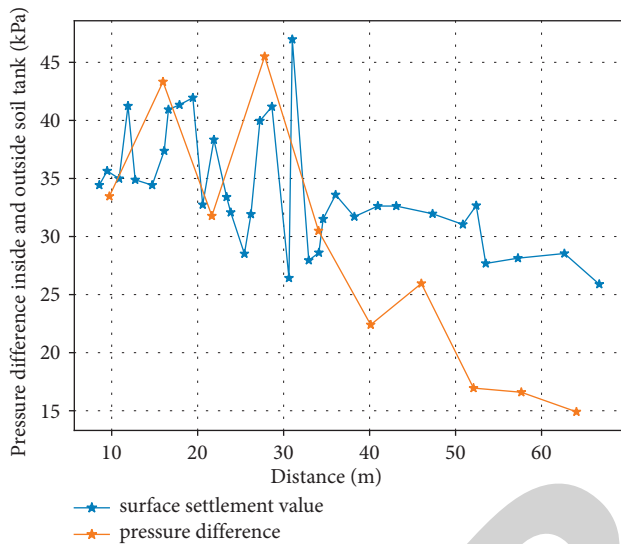


FIGURE 3: Relationship between surface settlement and pressure difference between soil and interior.

mud inlet and outlet conditions and then analyze and provide feedback in time to adjust the excavation parameters. The following is an example of the construction process from 101 to 200 rings through the shantytown sensitive area to analyze the regulation and control of tunneling parameters during the construction process.

In this project, the notch pressure is taken on the basis of the theoretical calculated value plus the corrected value of 0.02 MPa when the test section is dug, and Figure 4 shows the actual notch water pressure of the shield through the shantytown. The fluctuation of the water pressure at the incision should be controlled at -10 to 10 kPa to ensure the stability of the soil [10]. During the tunneling process, it was found that the mud and water pressure at the shield incision of the east line was greater than that of the west line, and the shield incision produced a larger uplift on the ground side, and the ground uplift in turn offset some of the stratigraphic losses during the construction process.

Shield initiation and advancement speed should be controlled within a reasonable range to ensure the normal excavation of the shield. Under normal conditions, tunneling speed of shield tunneling is $15\text{--}30$ mm/min [4]; the tunneling speed of shield machine in special strata should be controlled by $10\text{--}20$ mm/min. The speed of crossing the shantytown section should be controlled appropriately, but not too fast, which affects the follow-up speed of synchronous grouting and easily causes the cavity behind the

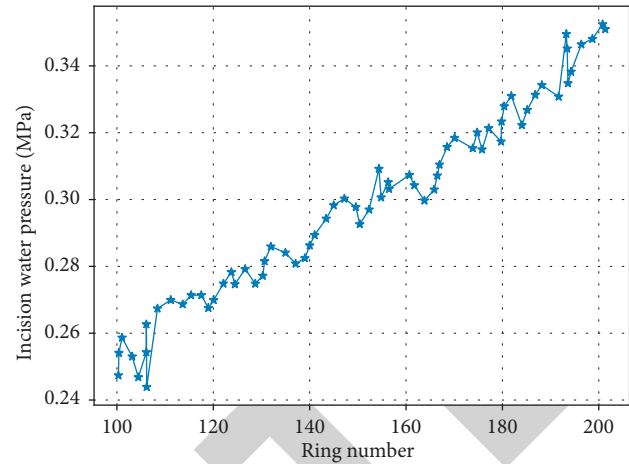


FIGURE 4: Actual notch water pressure.

pipe sheet wall and causes later settlement, so the matching of tunneling speed is essential for crossing the shantytown. Therefore, matching speed of tunneling is crucial for crossing the shantytown.

Referring to the settlement control experience of Nanjing Crossing Subway Line 10 project through the powder sand layer, the speed of the shield under penetration section is controlled at $0.8\text{--}0.9$ r/min, and the speed can be increased appropriately but not more than 1.0 r/min when the propulsion speed increases and the cone entry degree exceeds 50 . Figure 5 shows the tunneling speed of the shield under the soil structure.

The tunneling process of this construction has appropriately increased the tunneling speed on the basis of ensuring uniformity and stability. By appropriately increasing the tunneling speed, the construction disturbance and stratigraphic loss can be reduced to a certain extent, thus reducing the ground settlement caused by the extrusion of the shield machine's subsequent trailer on the subterranean layer.

Cutter torque is controlled at $6\text{--}9$ MN·m when crossing the shantytown, and the larger the cone penetration in the same geology, the higher the torque. When the torque gradually increases significantly when the parameters such as taper entry and speed remain unchanged, it is necessary to consider whether the tool is seriously worn, and the tool wear directly causes the torque to increase significantly, so it is necessary to stop and check the tool. Figure 6 shows the torque diagram of the shield under the shantytown construction tool.

In the soft soil layer, the rotating speed of the cutter should not be too large; otherwise, it will cause the tool to wear out faster, and the high rotating speed will easily cause greater disturbance to the soft soil layer, and the rotating speed of the shield through the shantytown section is $0.8\text{--}1.2$ r/min, and the rotating speed can be increased appropriately when the advancing speed increases and the cone entry degree exceeds 50 , but not more than 1.2 r/min. It is found that the size of the excavated clouds is reduced, thus effectively avoiding the phenomenon that the shield machine is prone to clogging the tunneling mud pipe in clay strata.

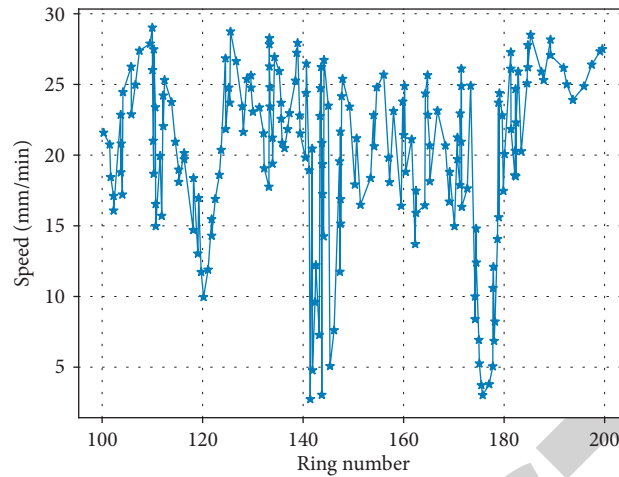


FIGURE 5: Scavenging speed of construction of 101 ~ 200 rings under shield.

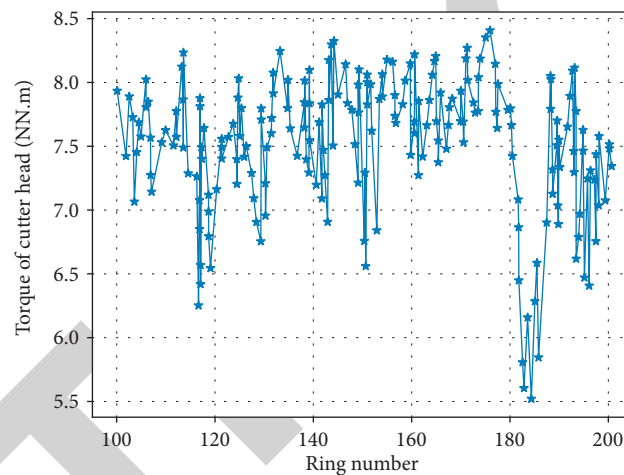


FIGURE 6: Shield torque diagram.

6. Synchronized Grouting Control

Table 2 shows the grouting rate and performance parameters per cubic meter in the excavation of the shanty section by the shield

The grouting pressure should ensure that the circumferential voids are filled and the structure of the cuttings is not damaged and deformed [11]. According to experience, the calculation of the grouting pressure can be calculated by slightly more than 0.05 MPa of the cutting pressure, and the specific setting is 0.3~0.5 MPa [32].

The gap formed by the excavation of the shield machine and the outer diameter of the pipe is $17.41 \text{ m}^3/\text{ring}$, and the filling factor of the shield through the shantytown section is 130%~180% according to experience of similar projects. At the same time, considering the shallow overburden and the influence of shielding uplift, the amount of grouting at the top and bottom of the shield should be controlled by 2:1. When the curve section is dug, a comparison of the grouting volume inside and outside of the curve section shows that the outside should be slightly larger than the inside.

7. Post-Crossing Settlement Monitoring and Pressure Fill Grouting

The shield machine continues to carry out monitoring and measurement work within a certain period of time after crossing, and Figures 7 and 8 show the monitoring curves of surface settlement and building settlement during shield crossing, respectively.

Analysis of Figures 7 and 8 shows that the shield tunnel shows a settlement trend before the palm face approaches the incision, and due to the high grouting pressure and easy disturbance of the fine sand stratum, there is a rapid uplift of the stratum in each monitoring section during the shield tunneling process, and then the soil starts to creep, consolidate, and shrink at the same time as the shield tail grouting, and the surface settlement value decreases.

The surface settlement and building settlement during the shield crossing should be controlled within the design range. At the same time, after fully considering the requirements for environmental protection and tunnel stability, if the radar grouting behind the wall detects or finds

TABLE 2: Synchronous grouting ratio and performance parameters.

P. O42.5 cement (kg)	Grade II fly ash (kg)	Bentonite (kg)	Fine sand (kg)	Water-reducing agent (kg)	Drinking water (kg)	Setting time (h)	Decantation rate (%)	Stone rate (%)	Consistency (cm)	Final setting strength (MPa)
200	200	80	1250	5.3	470	6	3	97	11	3.9

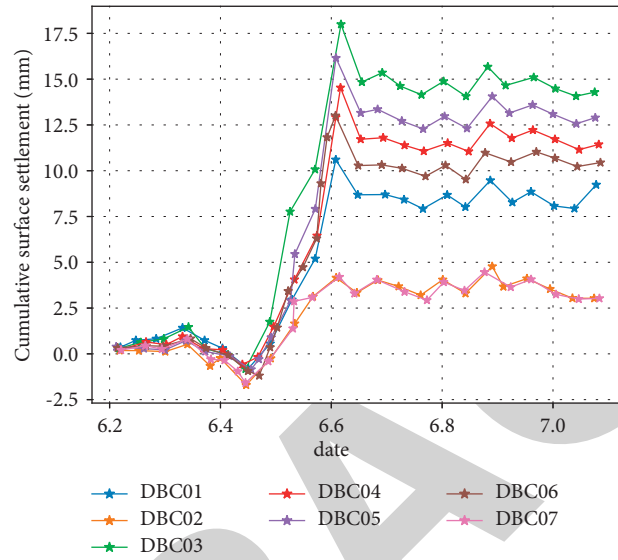


FIGURE 7: DC03 point longitudinal settlement cumulative duration curve.

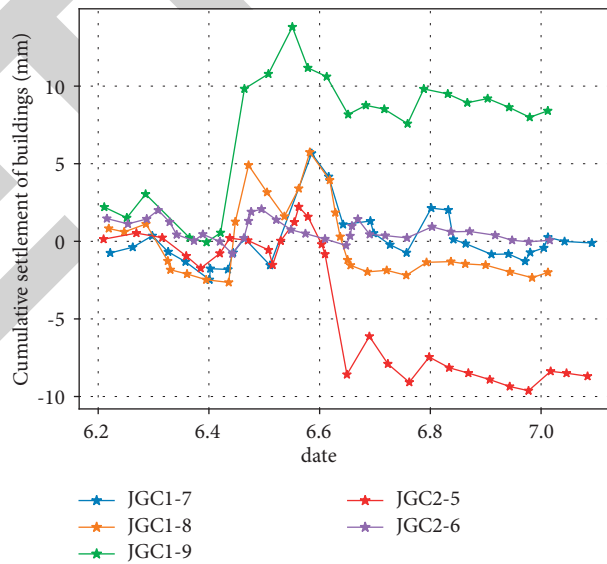


FIGURE 8: Building cumulative settlement distribution.

that the ground settlement still has a large trend of change or the local strata are soft, grouting should be replenished in time. Also, the stripping state caused by shield thrust should be filled in time to improve the water stopping effect.

8. Conclusions

In this study, the settlement control technology is studied and control measures are proposed for the construction of mega-diameter mud and water shield through shantytowns, using Wuhan Metro Line 8 Huangpu Road Station-Xujiapang Road Station interval shield through old shantytowns as the engineering background. The conclusions are as follows.

Before the shield starts, safety assessment of existing buildings (structures) should be carried out and safety control measures should be formulated; meanwhile, the tunneling parameters should be optimized through the tunneling test section.

In the process of shield tunneling, reasonable setting of each parameter should be combined with the geological condition and surrounding environment. The fluctuation value of notch pressure should be controlled by $-10\sim 10$ kPa to ensure the stability of soil body, and the starting digging speed should not be too fast and should be controlled at $10\sim 20$ mm/min in special sections.

The maximum uplift of the foundation of the old structure in the upper shantytown during the excavation of the oversized shield is 18 mm, and the overall settlement of the structure is controlled within 18 mm.

Data Availability

The dataset used in this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] S. Zhang, H. Liu, J. He, S. Han, and X. Du, "A deep bi-directional prediction model for live streaming recommendation," *Information Processing & Management*, vol. 58, no. 2, p. 102453, 2021.
- [2] M. Yang, C. H. Yeh, and Y. Zhou, "Design of an always-on deep neural network-based $1\text{-}\mu\text{W}$ voice activity detector aided with a customized software model for analog feature extraction," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 99, pp. 1–14, 2019, Joao P. C.
- [3] K. Merceron and A. Prouteau, "Évaluation de la cognition sociale en langue française chez l'adulte: outils disponibles et recommandations de bonne pratique clinique," *L'Évolution Psychiatrique*, vol. 78, no. 1, pp. 53–70, 2013.
- [4] S. A. Saeedinia, M. R. Jahed-Motlagh, A. Tafakhori, and N. Kasabov, "Design of MRI structured spiking neural networks and learning algorithms for personalized modelling, analysis, and prediction of EEG signals," *Scientific Reports*, vol. 11, no. 1, Article ID 12064, 2021.
- [5] S. Wang, B. Mo, and J. Zhao, "Deep neural networks for choice analysis: architecture design with alternative-specific utility functions-ScienceDirect," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 234–251, 2020.
- [6] N. Charaniya and S. Dudul, "Design of neural network models for daily rainfall prediction," *International Journal of Computer Application*, vol. 61, no. 14, pp. 23–27, 2013.
- [7] H. Nassar and H. Al-Mahdi, "Design and analysis of a TDMA call assignment scheme for cellular networks," *Computer Communications*, vol. 32, no. 7–10, pp. 1200–1206, 2009.
- [8] S. Rho, S. Song, Y. Nam, E. Hwang, and M. Kim, "Implementing situation-aware and user-adaptive music recommendation service in semantic web and real-time multimedia computing environment," *Multimedia Tools and Applications*, vol. 65, no. 2, pp. 259–282, 2013.
- [9] R. Reed, A. Bilgic, and R. Gotzhein, "Model-Driven development of time-critical protocols with SDL-MDD," *Part of the Lecture Notes in Computer Science*, vol. 579, pp. 34–52.
- [10] S. Pramanik, R. Haldar, and A. Kumar, "Deep learning driven venue recommender for event-based social networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 99, p. 1, 2019.
- [11] W. Yan, W. Dong, and M. Cao, "Deep auto encoder model with convolutional text networks for video recommendation," *IEEE Access*, vol. 7, no. 99, p. 1, 2019.
- [12] C. Weihs, U. Ligges, F. Mörchen, and D. Müllensiefen, "Classification in music research," *Advances in Data Analysis and Classification*, vol. 1, no. 3, pp. 255–291, 2007.
- [13] L. Wang, C. Zhang, Q. Chen et al., "A communication strategy of proactive nodes based on loop theorem in wireless sensor networks," in *Proceedings of the 2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 160–167, IEEE, Wanzhou, China, November 2018.
- [14] K. J. Kim and N. Baek: A Deep Learning Approach to Mining the Relationship of Depression Symptoms and Treatments for Prediction and Recommendation, 2019.
- [15] J. Xie, F. Zhu, M. Huang, N. Xiong, S. Huang, and W. Xiong, "Unsupervised learning of paragraph embeddings for context-aware recommendation," *IEEE Access*, vol. 7, pp. 43100–43109, 2019.
- [16] M. Zaslow, T. Halle, L. Martin et al., "Child outcome measures in the study of child care quality," *Evaluation Review*, vol. 30, no. 5, pp. 577–610, 2006.
- [17] L. Qiang, W. Shu, and W. Liang, "DeepStyle: learning user preferences for visual recommendation," in *Proceedings of the 40th International ACM SIGIR Conference ACM*, pp. 841–844, Shinjuku Tokyo Japan, August 2017.
- [18] S. Zhou, G.-L. Ye, L. Han, and W. Jian-Hua, "Key construction technologies for large river-crossing slurry shield tunnel: case study," *Journal of Aerospace Engineering*, vol. 34, no. 2, Article ID 04020118, 2021.
- [19] J. Liu, T. Qi, and Z. Wu, "Analysis of ground movement due to metro station driven with enlarging shield tunnels under building and its parameter sensitivity analysis," *Tunnelling and Underground Space Technology*, vol. 28, pp. 287–296, 2012.
- [20] H. Li, D. Zeng, L. Chen, Q. Chen, M. Wang, and C. Zhang, "Immune multipath reliable transmission with fault tolerance in wireless sensor networks," in *International Conference on Bio-Inspired Computing: Theories and Applications*, pp. 513–517, Springer, Singapore, October 2016.
- [21] J. Lai, H. Zhou, K. Wang et al., "Shield-driven induced ground surface and Ming Dynasty city wall settlement of Xi'an

Research Article

Risk Analysis of Textile Industry Foreign Investment Based on Deep Learning

Jingyi Liu¹ and Jiaolong Li² 

¹*School of Art and Design, South-Central University for Nationalities, Wuhan 430074, China*

²*Department of Mathematics and Quantitative Economics, School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China*

Correspondence should be addressed to Jiaolong Li; lijiaolong91@zuel.edu.cn

Received 6 December 2021; Accepted 25 December 2021; Published 10 January 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Jingyi Liu and Jiaolong 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 decline of China's economic growth rate and the uproar of antiglobalization, the textile industry, one of the business cards of China's globalization, is facing a huge impact. When the economic model is undergoing transformation, it is more important to prevent enterprises from falling into financial distress. So, the financial risk early warning is one of the important means to prevent enterprises from falling into financial distress. Aiming at the risk analysis of the textile industry's foreign investment, this paper proposes an analysis method based on deep learning. This method combines residual network (ResNet) and long short-term memory (LSTM) risk prediction model. This method first establishes a risk indicator system for the textile industry and then uses ResNet to complete deep feature extraction, which are further used for LSTM training and testing. The performance of the proposed method is tested based on part of the measured data, and the results show the effectiveness of the proposed method.

1. Introduction

With the slowdown of China's economic growth rate and changes in the growth model, the original crude corporate model has become unsustainable [1–4]. It has become a matter of course for the people to increase their consumption expectations based on the continuous improvement of living material standards. As a dual-intensive industry of labor and capital, the development of the textile industry plays a key role in the development of China's economy and guarantees the level of employment. Compared with other industries, the textile industry has the characteristics of long process, wide distribution, large number of employees, intensive capital, obvious geographical concentration, and high export ratio. These characteristics have brought the complexity and high risk of China's textile industry. Since the Sino-US trade friction in 2018, the textile industry as the business card of China's manufacturing industry has been frequently hit. The list of

additional tariffs imposed by the United States includes textile-related products, which has had a considerable impact on the Chinese textile industry. As an important part of China's foreign investment, scientifically conducting investment risk analysis in the textile industry plays an important role in improving economic benefits.

After years of development in the study of economic forecasting methods, a large number of forecasting models have emerged [5–15]. These models are divided into two categories: one is based on time series mainly including moving average and trend extrapolation and the other is based on causality, mainly including regression analysis, Markov prediction, and artificial neural network. In addition, the development of deep learning makes the fitting of complex systems more accurate. There are many studies on macroeconomic forecasts at home and abroad. Traditional economic forecasting methods, such as ARIMA and linear regression, have great limitations. ARIMA requires time series data to be stable. The linear regression is poor in fitting

complex nonlinear systems. Aiming at the complex nonlinear relationship in macroeconomic forecasting, the neural network model with strong ability to fit nonlinear systems has become a hot spot in the research of macroeconomic forecasting all over the world. There are many studies on the use of BPNN to establish prediction models. Literature [6] built a transmission line project cost prediction model based on BPNN, which could accurately estimate the project cost using a small number of samples. Therefore, it was suitable for comparing the pros and cons of the project in the early stage. Literature [9] combined with the Radial Basis Function (RBF) neural network and unbiased gray model (GM) established a gray RBF neural network prediction model. Through the predictive analysis of fiscal revenue data, it is found that training with this model not only has a fast convergence rate but also strong ability and high model accuracy. Literature [13] proposed a hybrid radial basis neural network, which integrated ridge regression, regression numbers, and radial basis neural networks. It was proved by experiments to predict the daily average trend of stock indexes. This network had a good effect when there are complex nonlinear relations between variables and mutual dependence. After years of research and development, artificial neural networks and their various improved models still cannot completely get rid of the defect that they tend to fall into local minimums and cannot reflect the timing relationship between samples [16–24]. However, this timing relationship is common in the economic field, and predictive analysis is of great help. At the same time, long short-term memory (LSTM) in deep learning shows an excellent ability to process time series data in the field of prediction. Literature [18] combined the trend data of steel transaction prices in recent years and trained the LSTM model. This work was compared and analyzed with the support vector regression model and found that the LSTM neural network can more accurately predict the price trend of steel. Literature [20] used deep learning for several years of the Gross Domestic Product (GDP) data in China to establish a prediction model, and the results show that the prediction accuracy based on deep learning was significantly higher than that of ARMA, LR, and exponential regression. Literature [23] aimed at the random and nonlinear characteristics of traffic flow using LSTM and Gated Recurrent Unit (GRU) neural network methods to predict short-term traffic flow. The experiments had proved that the deep learning methods based on recurrent neural network LSTM and GRU performed much better than the ARIMA models and some other methods [25–29].

Based on the existing methods, this paper proposes a textile industry foreign investment risk prediction method by combination of LSTM and ResNet [30–33]. First, an indicator system is established for investment risks in the textile industry, and feature vectors are constructed to describe the risk levels in the current state. On this basis, ResNet is used to perform further feature learning on the constructed index feature quantity to obtain deep features with stronger descriptive ability. As a reliable deep learning prediction method, LSTM predicts the characteristics learned by ResNet and obtains the current risk

representation. So, the current risk status of foreign investment in the textile industry can be judged. In the experiments, the proposed method is tested and verified with part of the data obtained publicly, and the results showed the effectiveness of the proposed method.

2. Basic Theory

2.1. LSTM. LSTM network is a model of memory cell network structure proposed by Hochreiter to solve the gradient explosion and gradient disappearance phenomenon that occurs when the recurrent neural network is processing relatively long time series data. It is introduced into the cell on the basis of Recurrent Neural Network (RNN). The threshold structure for judging whether the information meets the requirements is used to control the accumulation speed of information—input gate, forget gate, and output gate—so as to use this structure to memorize and update new information and solve the problem of long-term dependence. As shown in Figure 1, each LSTM neuron is composed of cell states, namely, long-term state c_t and short-term state h_t , input gate i_t , forget gate f_t , and output gate o_t .

The so-called cell state is a container for storing information. Through the process control of input gate, forget gate, and output gate, the information in the container is gradually increased, decreased, changed, and output. In each neural unit, the cell state undergoes the forgetting process of the forget gate, the input process of the input gate, and the process of outputting information to the output gate. The input gate is to copy and process the input information of the current neural unit. It consists of two parts: the sigmoid function independently chooses which information to update and the tanh function adds the constructed brand new vector to the current cell state to construct a new state. The realization formula is as follows:

$$\begin{cases} i_t = \sigma(w_{xi}x_t + w_{hi}h_{t-1} + b_i), \\ \tilde{c}_t = \phi(w_{xc}x_t + w_{hc}h_{t-1} + b_c). \end{cases} \quad (1)$$

The main function of the forget gate is to determine which information needs to be discarded in the current state. $f_t = 1$ means that the information is completely retained; $f_t = 0$ means that the information is completely discarded. The realization formula of f_t is as follows:

$$f_t = \sigma(w_{xf}x_t + w_{hf}h_{t-1} + b_f). \quad (2)$$

The output gate mainly controls the output information of the current hidden state. The realization formula is as follows:

$$\begin{cases} h_t = o_t \tanh(c_t), \\ c_t = i_t \tilde{c}_t + f_t c_{t-1}, \\ o_t = \sigma(w_{xo}x_t + w_{ho}h_{t-1} + b_o). \end{cases} \quad (3)$$

In equation (3), h_{t-1} is the output at the time $t - 1$; w_{xi} , w_{xc} , and w_{xo} are the matrix weight parameters for the input vector x_t the at the time t , respectively; w_{hi} , w_{hc} , and w_{ho} represent the weight matrix parameters of the hidden layer

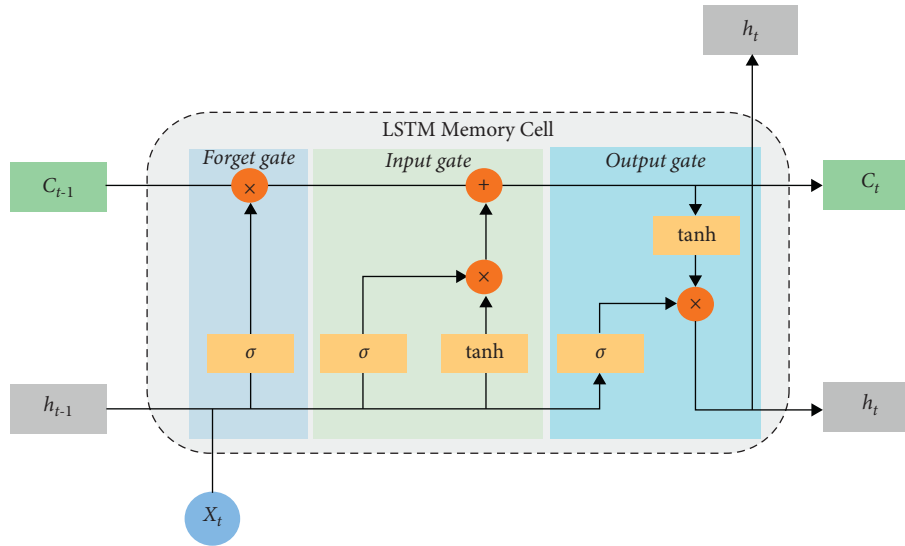


FIGURE 1: Basic structure of LSTM.

vector h_{t-1} at the time $t - 1$; and $b_i, b_c, b_f,$ and b_o represent the bias vector parameters.

Finally, the back propagation through time (BPTT) algorithm can be used to optimize the parameters of the LSTM model to obtain a reliable prediction model.

2.2. ResNet. As the number of deep neural network layers continues to increase, the learning ability of the network becomes stronger and stronger. However, the convergence speed of the relative network will slow down, and the gradient will disappear during the propagation process, making it impossible to effectively adjust the weights of the previous network layers. In the traditional convolutional neural networks (CNNs), except for the first layer, the input of each layer is derived from the output of the previous layer. The ResNet adopts a skip structure so that the deep residual network can directly cross the middle layers. The parameters are passed to the subsequent layers, which reduces the complexity of the network, solves the degradation problem of the deep-level network, and promotes the improvement of network performance.

The network structure of the residual neural network is shown in Figure 2. The network draws on the idea of cross-layer connection of high-speed networks. In the residual unit structure shown in Figure 2, x is the input of the network; $H(x)$ is the optimal solution mapping; and $F(x)$ represents the residual term and directly passes the input x to the output as the initial result to get the output $H(x) = F(x) + x$. When $F(x) = 0$, $H(x) = x$ is an equal mapping. The training goal of ResNet is to make the residual term $F(x) = H(x) - x$ close to zero. It is easier to let $F(x) = 0$ than $H(x) = x$. The updated parameters using $F(x) = 0$ can converge faster. Compared with the network model that does not adopt the skip structure, the residual network of this structure has clearer input data and can retain the accuracy of the data to the greatest extent.

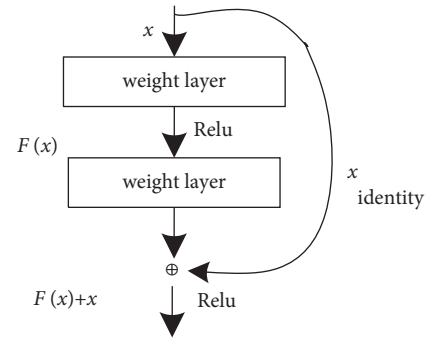


FIGURE 2: Illustration of residual network unit.

3. Risk Assessment Method

3.1. Index System. The purpose of financial early warning indicators is to detect the possible financial risks of the enterprise in advance and to rate the risks. Therefore, when selecting indicators, you should choose indicators with the nature of capability evaluation and risk evaluation. Therefore, based on relevant domestic and foreign research results, this article selects the total return on net assets, net sales interest rate, net operating cash flow/debt, and operating activities from five aspects: profitability, operating ability, development ability, debt solvency, and cash flow ability. A total of 10 financial indicators including net cash flow growth rate, net asset growth rate, total asset growth rate, accounts receivable turnover rate, current asset turnover rate, asset-liability ratio, and asset-liability ratio are statistically analyzed. Based on the data indicators of the textile industry's foreign investment, the feature vectors of these 10 indicators can be constructed accordingly, which can be used for subsequent risk prediction and evaluation.

3.2. Evaluation Process. This paper builds a ResNet-LSTM-based textile industry foreign investment risk analysis model based on the previous discussion. The proposed method

makes full use of the advantages of ResNet and better retains the attributes of the original data so as to effectively solve the problem of extracting the characteristics of financial risk data in the textile industry. Using the advantages of LSTM in data prediction, the accuracy of risk prediction is improved through ResNet's deep features. The basic flow chart of the method in this paper is shown in Figure 3, and the specific steps are described as follows:

- (1) Based on the historical data set of the textile industry's foreign investment, the feature vector is constructed according to the established index system.
- (2) ResNet is used for deep feature extraction to further optimize the feature vector.
- (3) In the training phase, the deep features extracted by the optimized ResNet are input into the LSTM network, and the unique characteristics of the memory unit structure of the LSTM network are used to establish a prediction model.
- (4) In the testing phase, the investment financial data set of the current data is input to the established LSTM prediction model, and the current risk assessment result is output.

3.3. Evaluation Index. Assuming the predicted value as $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ and the true value as $y = \{y_1, y_2, \dots, y_n\}$, the root mean square error (RMSE), mean absolute percentage error (MAPE), and mean accuracy (MA) are used as the evaluation indicators of the prediction model. The definitions of the three indexes are as follows:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \\ \text{MAPE} &= \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\%, \\ \text{MA} &= 1 - \text{MAPE}. \end{aligned} \quad (4)$$

In the formula, the smaller the RMSE value, the closer the predicted value to the true value, the higher the accuracy of the prediction. MAPE and MA evaluate the predictive power of the model. The smaller the MAPE and the larger the MA, the better the prediction effect of the model.

4. Experiment and Analysis

4.1. Data Set and Comparison Method. In order to test the performance of the proposed method, this paper obtained 380 historical financial data of China's textile industry foreign investment through public channels. According to the financial risk measurement index system in the previous

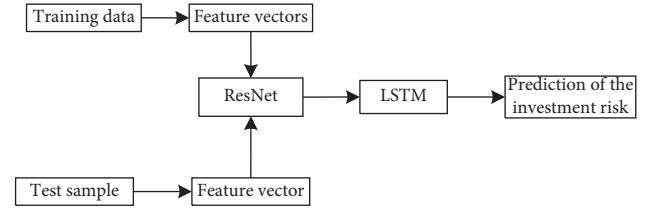


FIGURE 3: Prediction of investment risk based on LSTM and ResNet.

works, the corresponding feature vectors are constructed, respectively. Then, the method of combining ResNet and LSTM is used for risk prediction analysis, and RMSE and MA indicators are used to evaluate the performance of the method. While testing the method in this paper, some existing methods are selected for comparative analysis, including the method based on support vector machine (SVM), the method based on BP network, and the method based on LSTM.

5. Results and Analysis

The selected investment data are tested using the proposed method, and the performance of the method in this paper and the comparison methods is as shown in Table 1. According to the results, the performance of the method in this paper is the best among the four types of methods, reflecting its performance advantages. Among the three types of comparison methods, LSTM has the best performance, indicating its significant advantages in data prediction. BP network also has certain advantages over SVM, which shows the advantages of neural network. In particular, comparing the proposed method with LSTM shows that this paper further introduces ResNet for deep feature learning, which further improves the final prediction performance.

Therefore, by using the comprehensive evaluation of RMSE and MAPE two indicators, the proposed method has advantages in predicting the risk of foreign investment in the textile industry.

In the actual process, due to the influence of market bands and other political and economic factors, certain errors may occur in the prediction model. To this end, this paper applies a certain degree of noise to the experimental data to reflect the volatility of economic data. On this basis, MA was used as the basic evaluation index to test the performance trend of various methods, and the results are shown in Figure 4. It can be seen from the figure that the performance of various methods is degraded to a certain extent due to the influence of noise. In comparison, the method in this paper can maintain the best prediction performance under different noise interference conditions, thereby further improving its performance advantages.

TABLE 1: Comparison of prediction performance of different methods.

	RMSE	MA (%)
Proposed	0.14	97.3
SVM	0.21	95.6
BP	0.18	96.1
LSTM	0.15	96.8

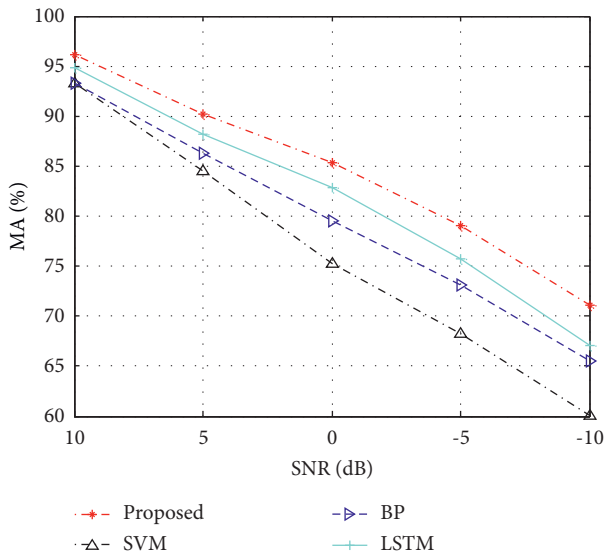


FIGURE 4: Performance of different methods under noises.

6. Conclusion

Scientific analysis of the risk of foreign investment in the textile industry is conducive to identifying the investment direction and investment market and obtaining the greatest rate of return. Aiming at the problem of predicting the risk of foreign investment in the textile industry, this paper comprehensively uses two deep learning models, ResNet and LSTM. ResNet is used for deep feature learning to further optimize the index system of risk prediction. As an accurate and efficient prediction model, LSTM predicts the deep features learned by ResNet and obtains risk prediction results. Part of the measured data is used to verify and analyze the proposed method, and the result proves the performance advantage of the proposed method. Compared with the method directly using LSTM, the proposed can bring performance improvement owing to the merits of ResNet. In the future, more suitable deep learning models can be further developed to improve the accuracy of the prediction model.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Lizhen Du, Lianqing Yu, and Rao Cheng, "The construction research on Rapid-Response Eco-Supply Chain of the textile industry based on the circular economy," in *Proceedings of the 2010 International Conference on E-Health Networking Digital Ecosystems and Technologies (EDT)*, pp. 248–251, Shenzhen, China, April 2010.
- [2] J. Liu, J. Jiang, and Q. Zhang, "Research on geo-relationship network and competing (or mutually beneficial) relationship network of Chinese textile enterprises," in *Proceedings of the 2017 29th Chinese Control And Decision Conference (CCDC)*, pp. 6462–6467, Chongqing, China, May 2017.
- [3] Z. Haiyan, "The Improvement of the Industry Structure of Chinese Textile Industry on the Basis of Independent Innovations," in *Proceedings of the 2009 First International Workshop on Education Technology and Computer Science*, pp. 422–424, Wuhan, Hubei, China, May 2009.
- [4] K. Wang, X. Zeng, L. Koehl, X. Tao, and Y. Chen, "Statistical based approach for uncertainty analysis in life cycle assessment: a case study in textile industry," in *Proceedings of the 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–4, LA, USA, June 2019.
- [5] W. Jiang, L. Deng, L. Chen, J. Wu, and J. Li, "Risk assessment and validation of flood disaster based on fuzzy mathematics," *Progress in Natural Science*, vol. 19, no. 10, pp. 1419–1425, 2009.
- [6] Enlin Tang, "Research on Interest Rate Risk of Housing Mortgage Loan Based on Computer Simulation," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6035022, 6 pages, 2021.
- [7] Hui Chu, "An Empirical Analysis of Corporate Financial Management Risk Prediction Based on Associative Memory Neural Network," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 4383742, 11 pages, 2021.
- [8] Wenjuan Liu, "Enterprise Credit Risk Management Using Multicriteria Decision-Making," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6191167, 10 pages, 2021.
- [9] Boning Huang, Junkang Wei, Yuhong Tang, and Liu Chang, "Enterprise Risk Assessment Based on Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6049195, 6 pages, 2021.
- [10] A Jian Luo, B Xin Yan, and Ye Tian, "Unsupervised quadratic surface support vector machine with application to credit risk assessment," *European Journal of Operational Research*, vol. 280, pp. 1008–1017, 2019.
- [11] Mirko Moscatellia, Fabio Parlapianoa, Simone Narizzanob, and Gianluca Viggiano, "Corporate default forecasting with machine learning," *Expert Systems with Applications*, vol. 161, no. 6, pp. 1–12, Article ID 113567, 2020.
- [12] N. Arora and P. Deep Kaur, "A Bolasso based consistent feature selection enabled random forest classification algorithm: An application to credit risk assessment," *Applied Soft Computing Journal*, vol. 86, pp. 1–15, Article ID 105936, 2020.
- [13] Diego Paganoti Fonseca, Peter Fernandes Wankea, and Henrique Luiz Correa, "A two-stage fuzzy neural approach for credit risk assessment in a Brazilian credit card company," *Applied Soft Computing Journal*, vol. 92, pp. 1–13, 2020.
- [14] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [15] Yuelin Wang, Yihan Zhanga, Lua Yan, and Xinran Yua, "A Comparative Assessment of Credit Risk Model Based on Machine Learning—a case study of bank loan data [J]," *Procedia Computer Science*, vol. 174, pp. 141–149, 2020.

- [16] J Jin Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 3, pp. 299–310, 2005.
- [17] D. D. Wu, D. L. Olson, and C. Luo, "A Decision Support Approach for Accounts Receivable Risk Management," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 12, pp. 1624–1632, 2014.
- [18] J. W. Nowak, S. Sarkani, and T. A. Mazzuchi, "Risk Assessment for a National Renewable Energy Target Part II: Employing the Model," *IEEE Systems Journal*, vol. 10, no. 2, pp. 459–470, 2016.
- [19] N. Wan, L. Li, C. Ye, and B. Wang, "Risk Assessment in Intelligent Manufacturing Process: A Case Study of an Optical Cable Automatic Arranging Robot," *IEEE Access*, vol. 7, pp. 105892–105901, 2019.
- [20] Z. Qu, "Application of improved PCA in risks assessment technology of enterprise information security," in *Proceedings of the 2009 2nd International Conference on Power Electronics and Intelligent Transportation System (PEITS)*, pp. 58–61, Shenzhen, China, December 2009.
- [21] Zhandong Xu and G. Chi, "Bank-enterprise project risk assessment model based on the information entropy method," in *proceedings of the 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, pp. 998–1001, Zhengzhou, China, August 2011.
- [22] P.-B. Zhang and Z.-X. Yang, "A Novel AdaBoost Framework With Robust Threshold and Structural Optimization," *IEEE Transactions on Cybernetics*, vol. 48, no. 1, pp. 64–76, 2018.
- [23] S. Wu and H. Nagahashi, "Parameterized AdaBoost: Introducing a Parameter to Speed Up the Training of Real AdaBoost," *IEEE Signal Processing Letters*, vol. 21, no. 6, pp. 687–691, 2014.
- [24] Qu Qiao, Cheng Liu, and Xinzhong Bao, "E-Commerce Enterprise Supply Chain Financing Risk Assessment Based on Linked Data Mining and Edge Computing," *Mobile Information Systems*, vol. 2021, Article ID 9938325, 9 pages, 2021.
- [25] Tingqiang Chen, Qinghao Yang, Yutong Wang, and Suyang Wang, "Double-Layer Network Model of Bank-Enterprise Counterparty Credit Risk Contagion," *Complexity*, vol. 2020, Article ID 3690848, 25 pages, 2020.
- [26] Yichuan Shao, Xingjia Yao, Liwei Tian, and Hanning Chen, "A Multiswarm Optimizer for Distributed Decision Making in Virtual Enterprise Risk Management," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 904815, 24 pages, 2012.
- [27] Shanshan Dong and Liu Chang, "Sentiment Classification for Financial Texts Based on Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9524705, 9 pages, 2021.
- [28] Xianfu Wei, "A Method of Enterprise Financial Risk Analysis and Early Warning Based on Decision Tree Model," *Security and Communication Networks*, vol. 2021, Article ID 6950711, 9 pages, 2021.
- [29] Jun Wu, Chengbing Li, and Yueying Huo, "Safety Assessment of Dangerous Goods Transport Enterprise Based on the Relative Entropy Aggregation in Group Decision Making Model," *Computational Intelligence and Neuroscience*, vol. 2014, Article ID 571058, 7 pages, 2014.
- [30] Alessio Galassi, D José, Martín Guerrero, Eduardo Villamor, Carlos Monserrat, and María José Rupérez, "Risk Assessment of Hip Fracture Based on Machine Learning," *Applied Bionics and Biomechanics*, vol. 2020, Article ID 8880786, 13 pages, 2020.
- [31] Brian Williams, Brandon Allen, Zhen Hu et al., "Real-Time Fall Risk Assessment Using Functional Reach Test," *International Journal of Telemedicine and Applications*, vol. 2017, Article ID 2042974, 8 pages, 2017.
- [32] Y. Jianxing, C. Haicheng, W. Shibo, and F. Haizhao, "A Novel Risk Matrix Approach Based on Cloud Model for Risk Assessment Under Uncertainty," *IEEE Access*, vol. 9, pp. 27884–27896, 2021.
- [33] B. R. Greene, S. J. Redmond, and B. Caulfield, "Fall Risk Assessment Through Automatic Combination of Clinical Fall Risk Factors and Body-Worn Sensor Data," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 725–731, 2017.